

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-125199

(43)Date of publication of application : 26.04.2002

(51)Int.Cl.

H04N 5/93  
G11B 20/10  
H04N 5/765  
H04N 5/781  
H04N 5/85

(21)Application number : 2001-200220

(71)Applicant : TOSHIBA CORP

(22)Date of filing : 29.06.2001

(72)Inventor : HORI OSAMU

KANEKO TOSHIMITSU

MITA TAKESHI

YAMAMOTO KOJI

MASUKURA KOICHI

(30)Priority

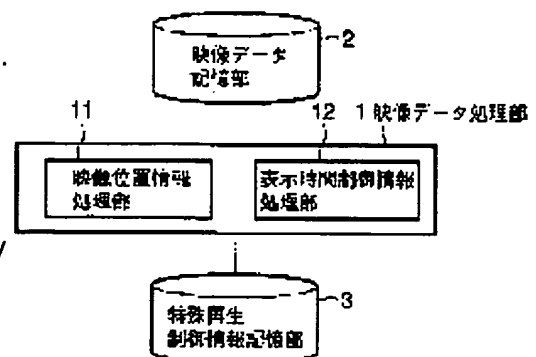
Priority number : 2000200220 Priority date : 30.06.2000 Priority country : JP

(54) FRAME INFORMATION DESCRIPTION METHOD, FRAME INFORMATION GENERATING DEVICE AND METHOD, VIDEO REPRODUCING DEVICE AND METHOD, AND RECORDING MEDIUM

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a special reproduction control information generator that generates special reproduction control information to allow a user to attain more effective special reproduction.

SOLUTION: In the special reproduction control information generator, a video position information processing section 11 reads object video data from a video data storage section 2, sequentially and selectively extracts part of frames used for special reproduction along frame sequences, generates video position information denoting a position of each extracted frame with respect to original video data and a display time



THIS PAGE BLANK (USPTO)

control information processing section 12 generates display time control information denoting a display time of each frame. Then the processing section 12 generates special reproduction control information where the video position information and the display time control information of each frame are arranged as frame information and a special reproduction control information storage section 3 stores the special reproduction control information.

---

## LEGAL STATUS

[Date of request for examination] 11.03.2005

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

THIS PAGE BLANK (USPTO)

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2002-125199

(P2002-125199A)

(43) 公開日 平成14年4月26日 (2002. 4. 26)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テーマコード* (参考)
H 0 4 N 5/93		G 1 1 B 20/10	3 2 1 Z 5 C 0 5 2
G 1 1 B 20/10	3 2 1	H 0 4 N 5/85	A 5 C 0 5 3
H 0 4 N 5/765		5/93	Z 5 D 0 4 4
5/781		5/781	5 1 0 G
5/85			

審査請求 未請求 請求項の数24 O L (全 31 頁)

(21) 出願番号 特願2001-200220 (P2001-200220)  
(22) 出願日 平成13年6月29日 (2001. 6. 29)  
(31) 優先権主張番号 特願2000-200220 (P2000-200220)  
(32) 優先日 平成12年6月30日 (2000. 6. 30)  
(33) 優先権主張国 日本 (J P)

(71) 出願人 000003078  
株式会社東芝  
東京都港区芝浦一丁目1番1号  
(72) 発明者 堀 修  
神奈川県川崎市幸区小向東芝町1番地 株  
式会社東芝研究開発センター内  
(72) 発明者 金子 敏充  
神奈川県川崎市幸区小向東芝町1番地 株  
式会社東芝研究開発センター内  
(74) 代理人 100058479  
弁理士 鈴江 武彦 (外6名)

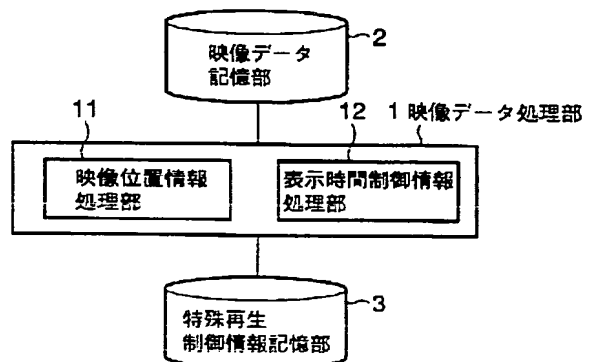
最終頁に続く

(54) 【発明の名称】 フレーム情報記述方法、フレーム情報生成装置及び方法、映像再生装置及び方法並びに記録媒体

(57) 【要約】

【課題】 利用者にとってより効果的な特殊再生を可能とする特殊再生制御情報を生成する特殊再生制御情報生成装置を提供すること。

【解決手段】 特殊再生制御情報生成装置では、映像データ記憶部2から対象となる映像データを読み込み、映像位置情報処理部11にて、特殊再生に供される一部のフレームをフレーム系列に沿って順次選択的に抽出し、抽出した各フレームの元映像データにおける位置を示す映像位置情報を作成し、表示時間制御情報処理部12にて、各フレームの表示時間を示す表示時間制御情報を作成する。そして、各フレームの映像位置情報と表示時間制御情報をフレーム情報として配列した特殊再生制御情報を作成し、これを特殊再生制御情報記憶部3に格納する。



## 【特許請求の範囲】

【請求項1】 元映像データの複数のフレームから抽出したフレームの該元映像データ中における位置を特定する第1情報を記述するステップと、  
該抽出フレームの表示時間に関する第2情報を記述するステップと、

を具備することを特徴とするフレーム情報記述方法。

【請求項2】 前記抽出フレームはフレーム群からなり、前記第1情報は該抽出フレーム群の該元映像データ中における位置を特定する情報を具備することを特徴とする請求項1記載のフレーム情報記述方法。

【請求項3】 該抽出フレームの重要度に関する第3情報を記述するステップをさらに具備することを特徴とする請求項1記載のフレーム情報記述方法。

【請求項4】 前記第1情報は前記元映像データから作成される該抽出フレームに対応する画像データファイルを特定する情報を具備することを特徴とする請求項1記載のフレーム情報記述方法。

【請求項5】 前記抽出フレームは前記元映像データのある時間的区間の複数フレームから抽出されたフレームを具備し、前記時間的区間を特定する第4情報を記述するステップをさらに具備することを特徴とする請求項1記載のフレーム情報記述方法。

【請求項6】 前記第1情報は前記元映像データから作成される該抽出フレームに対応する画像データファイルを特定する情報を具備することを特徴とする請求項5記載のフレーム情報記述方法。

【請求項7】 前記第2情報は再生時の画面変化量がほぼ一定となるような表示時間に関する情報を具備することを特徴とする請求項1記載のフレーム情報記述方法。

【請求項8】 前記抽出フレームの再生又は非再生を指示する第5情報を記述するステップをさらに具備することを特徴とする請求項1記載のフレーム情報記述方法。

【請求項9】 前記第1情報は前記抽出フレームの位置を示す情報、または前記元映像データから生成され、元映像データとは別に格納される画像データファイル中の該抽出フレームに対応する画像データの位置を示す情報を具備することを特徴とする請求項1記載のフレーム情報記述方法。

【請求項10】 前記抽出フレームを含む前記元映像データ以外のメディアデータに、該メディアデータの位置を示す情報と、該メディアデータの表示時間に関する情報を記述するステップをさらに具備することを特徴とする請求項1記載のフレーム情報記述方法。

【請求項11】 元映像データの複数のフレームから抽出したフレームに関するフレーム情報を格納するコンピュータ読取り可能な記録媒体において、前記フレーム情報は

該抽出フレームの前記元映像データ中における位置を特定する第1情報と、

該抽出フレームの表示時間に関する第2情報と、  
を具備することを特徴とする記録媒体。

【請求項12】 前記抽出フレームはフレーム群からなり、前記第1情報は該抽出フレーム群の該元映像データ中における位置を特定する情報を具備することを特徴とする請求項11記載の記録媒体。

【請求項13】 前記フレーム情報は該抽出フレームの重要度に関する第3情報をさらに具備することを特徴とする請求項11記載の記録媒体。

10 【請求項14】 前記第1情報は前記元映像データから作成される該抽出フレームに対応する画像データファイルを特定する情報を具備することを特徴とする請求項11記載の記録媒体。

【請求項15】 前記フレーム情報とともに、前記元映像データ、および前記元映像データから作成される前記抽出フレームに対応する画像データファイルを更に格納することを特徴とする請求項11記載の記録媒体。

20 【請求項16】 元映像データの複数のフレームから抽出したフレームの該元映像データ中における位置を特定する第1情報を記述する手段と、  
該抽出フレームの表示時間に関する第2情報を記述する手段と、

を具備することを特徴とするフレーム情報記述装置。

【請求項17】 元映像データの複数のフレームから抽出したフレームの該元映像データ中における位置を特定する第1情報を生成するステップと、  
該抽出フレームの表示時間に関する第2情報を生成するステップと、

30 を具備することを特徴とするフレーム情報生成方法。

【請求項18】 元映像データの複数のフレームから抽出したフレームの該元映像データ中における位置を特定する第1情報と、該抽出フレームの表示時間に関する第2情報とを参照する手段と、

前記第1情報に基づいて前記抽出フレームの元映像データを取得する手段と、

前記第2情報に基づいて前記抽出フレームの元映像データを再生する表示時間を決定する手段と、

取得された前記元映像データを決定された前記表示時間再生する手段と、

40 を具備することを特徴とする映像再生装置。

【請求項19】 元映像データの複数のフレームから抽出したフレームの該元映像データ中における位置を特定する第1情報と、該抽出フレームの表示時間に関する第2情報とを参照するステップと、

前記第1情報に基づいて前記抽出フレームの元映像データを取得するステップと、

前記第2情報に基づいて前記抽出フレームの元映像データを再生する表示時間を決定するステップと、

50 取得された前記元映像データを決定された前記表示時間再生するステップと、

を具備することを特徴とする映像再生方法。

【請求項 20】 映像再生プログラムを格納するコンピュータ読取り可能な記録媒体において、前記映像再生プログラムは

コンピュータに元映像データの複数のフレームから抽出したフレームの該元映像データ中における位置を特定する第 1 情報と、該抽出フレームの表示時間に関する第 2 情報とを参照させるプログラムコードと、

コンピュータに前記第 1 情報に基づいて前記抽出フレームの元映像データを取得させるプログラムコードと、コンピュータに前記第 2 情報に基づいて前記抽出フレームの元映像データを再生する表示時間を決定させるプログラムコードと、

コンピュータに取得された前記元映像データを決定された前記表示時間再生させるプログラムコードと、を具備することを特徴とする記録媒体。

【請求項 21】 オリジナル音データの複数のフレームから抽出した音フレームの該オリジナル音データ中における位置を特定する第 1 情報を記述するステップと、該抽出フレームの再生時間に関する第 2 情報を記述するステップと、

を具備することを特徴とするフレーム情報記述方法。

【請求項 22】 オリジナル音データの複数のフレームから抽出した音フレームに関するフレーム情報を格納するコンピュータ読取り可能な記録媒体において、前記フレーム情報は該抽出フレームの前記オリジナル音データ中における位置を特定する第 1 情報と、

該抽出フレームの再生時間に関する第 2 情報と、

を具備することを特徴とする記録媒体。

【請求項 23】 オリジナルテキストデータの複数のフレームから抽出したテキストフレームの位置を特定する第 1 情報を記述するステップと、

該抽出フレームの表示時間に関する第 2 情報を記述するステップと、

を具備することを特徴とするフレーム情報記述方法。

【請求項 24】 オリジナルテキストデータの複数のフレームから抽出したテキストフレームに関するフレーム情報を格納するコンピュータ読取り可能な記録媒体において、前記フレーム情報は該抽出フレームの前記オリジナルテキストデータ中における位置を特定する第 1 情報と、

該抽出フレームの表示時間に関する第 2 情報と、

を具備することを特徴とする記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、デジタルコンテンツ、例えば映像コンテンツを特殊再生するためのフレーム情報記述方法、フレーム情報生成装置及び方法、映像再生装置及び方法並びに記録媒体に関する。

【0002】

【従来の技術】近年、動画像をデジタル映像として圧縮し、DVDやHDDに代表されるディスクメディアに格納することによって、映像をランダム再生できる環境が整った。この環境では、あらかじめ定められた場所から、待ち時間がほとんどない状態で途中再生が可能である。従来のテープメディアと同様に 2～4 倍速の早回しおよび逆回しも可能である。

【0003】

【発明が解決しようとする課題】しかしながら、映像は長時間の場合が多く、2～4 倍速の再生でも全ての内容を視聴するには、時間が十分短縮されない問題があった。倍速再生を大きくすると、画面の変化が視聴能力以上に大きくなって内容把握が困難であったり、本来内容を把握するにはあまり重要でない部分においても、倍速再生が行われたりして無駄な場合がある。

【0004】このような問題は映像コンテンツに限らず、音・テキストからなるコンテンツ、あるいはマルチメディアコンテンツにも当てはまる。

【0005】本発明の目的は、利用者にとってより効果的な特殊再生を可能とするフレーム情報記述方法、フレーム情報生成装置及び方法、映像再生装置及び方法並びに記録媒体に関する。

【0006】

【課題を解決するための手段】上記した課題を解決し目的を達成するために、本発明は以下に示す手段を用いている。

【0007】本発明のフレーム情報記述方法は、元映像データの複数のフレームから抽出したフレームの該元映像データ中における位置を特定する第 1 情報を記述するステップと、該抽出フレームの表示時間に関する第 2 情報を記述するステップとを具備することを特徴とする。

【0008】前記抽出フレームはフレーム群からなり、前記第 1 情報は該抽出フレーム群の該元映像データ中における位置を特定する情報を具備してもよい。

【0009】該抽出フレームの重要度に関する第 3 情報を記述するステップをさらに具備してもよい。

【0010】前記第 1 情報は前記元映像データから作成される該抽出フレームに対応する画像データファイルを特定する情報を具備してもよい。

【0011】前記抽出フレームは前記元映像データのある時間的区間の複数フレームから抽出されたフレームを具備し、前記時間的区間を特定する第 4 情報を記述するステップをさらに具備してもよい。

【0012】前記第 1 情報は前記元映像データから作成される該抽出フレームに対応する画像データファイルを特定する情報を具備してもよい。

【0013】前記第 2 情報は再生時の画面変化量がほぼ一定となるような表示時間に関する情報を具備してもよい。

【0014】前記抽出フレームの再生又は非再生を指示

する第5情報を記述するステップをさらに具備してもよい。

【0015】前記第1情報は前記抽出フレームの位置を示す情報、または前記元映像データから生成され、元映像データとは別に格納される画像データファイル中の該抽出フレームに対応する画像データの位置を示す情報を具備してもよい。

【0016】前記抽出フレームを含む前記元映像データ以外のメディアデータに、該メディアデータの位置を示す情報と、該メディアデータの表示時間に関する情報を記述するステップをさらに具備してもよい。

【0017】本発明の元映像データの複数のフレームから抽出したフレームに関するフレーム情報を格納するコンピュータ読取り可能な記録媒体において、前記フレーム情報は該抽出フレームの前記元映像データ中における位置を特定する第1情報と、該抽出フレームの表示時間に関する第2情報とを具備する。

【0018】前記抽出フレームはフレーム群からなり、前記第1情報は該抽出フレーム群の該元映像データ中における位置を特定する情報を具備してもよい。

【0019】前記フレーム情報は該抽出フレームの重要度に関する第3情報をさらに具備してもよい。

【0020】前記第1情報は前記元映像データから作成される該抽出フレームに対応する画像データファイルを特定する情報を具備してもよい。

【0021】前記フレーム情報とともに、前記元映像データ、および前記元映像データから作成される前記抽出フレームに対応する画像データファイルを更に格納してもよい。

【0022】本発明のフレーム情報記述装置は元映像データの複数のフレームから抽出したフレームの該元映像データ中における位置を特定する第1情報を記述する手段と、該抽出フレームの表示時間に関する第2情報を記述する手段とを具備することを特徴とする。

【0023】本発明のフレーム情報生成方法は元映像データの複数のフレームから抽出したフレームの該元映像データ中における位置を特定する第1情報を生成するステップと、該抽出フレームの表示時間に関する第2情報を生成するステップとを具備することを特徴とする。

【0024】本発明の映像再生装置は元映像データの複数のフレームから抽出したフレームの該元映像データ中における位置を特定する第1情報と、該抽出フレームの表示時間に関する第2情報とを参照する手段と、前記第1情報に基づいて前記抽出フレームの元映像データを取得する手段と、前記第2情報に基づいて前記抽出フレームの元映像データを再生する表示時間を決定する手段と、取得された前記元映像データを決定された前記表示時間再生する手段とを具備することを特徴とする。

【0025】本発明の映像再生方法は元映像データの複数のフレームから抽出したフレームの該元映像データ中

における位置を特定する第1情報と、該抽出フレームの表示時間に関する第2情報とを参照するステップと、前記第1情報に基づいて前記抽出フレームの元映像データを取得するステップと、前記第2情報に基づいて前記抽出フレームの元映像データを再生する表示時間を決定するステップと、取得された前記元映像データを決定された前記表示時間再生するステップとを具備することを特徴とする。

【0026】本発明の映像再生プログラムを格納するコンピュータ読取り可能な記録媒体において、前記映像再生プログラムはコンピュータに元映像データの複数のフレームから抽出したフレームの該元映像データ中における位置を特定する第1情報と、該抽出フレームの表示時間に関する第2情報とを参照させるプログラムコードと、コンピュータに前記第1情報に基づいて前記抽出フレームの元映像データを取得させるプログラムコードと、コンピュータに前記第2情報に基づいて前記抽出フレームの元映像データを再生する表示時間を決定させるプログラムコードと、コンピュータに取得された前記元映像データを決定された前記表示時間再生させるプログラムコードとを具備することを特徴とする。

【0027】本発明のフレーム情報記述方法はオリジナル音データの複数のフレームから抽出した音フレームの該オリジナル音データ中における位置を特定する第1情報を記述するステップと、該抽出フレームの再生時間に関する第2情報を記述するステップとを具備することを特徴とする。

【0028】本発明のオリジナル音データの複数のフレームから抽出した音フレームに関するフレーム情報を格納するコンピュータ読取り可能な記録媒体において、前記フレーム情報は該抽出フレームの前記オリジナル音データ中における位置を特定する第1情報と、該抽出フレームの再生時間に関する第2情報とを具備することを特徴とする。

【0029】本発明のフレーム情報記述方法はオリジナルテキストデータの複数のフレームから抽出したテキストフレームの位置を特定する第1情報を記述するステップと、該抽出フレームの表示時間に関する第2情報を記述するステップとを具備することを特徴とする。

【0030】本発明のオリジナルテキストデータの複数のフレームから抽出したテキストフレームに関するフレーム情報を格納するコンピュータ読取り可能な記録媒体において、前記フレーム情報は該抽出フレームの前記オリジナルテキストデータ中における位置を特定する第1情報と、該抽出フレームの表示時間に関する第2情報とを具備することを特徴とする。

【0031】本発明によれば、利用者にとってより効果的な特殊再生を可能とするフレーム情報記述方法、フレーム情報生成装置及び方法、映像再生装置及び方法並びに記録媒体が提供される。



【0032】

【発明の実施の形態】以下、図面を参照しながら発明の実施の形態を説明する。

【0033】本発明は全てのデジタルコンテンツに関するが、実施形態として映像データを有する映像コンテンツの再生を説明する。映像データは、動画像を構成する映像フレームの集合（映像フレーム群）からなっているものとする。

【0034】最初に、本実施形態において重要な役割を果たす特殊再生制御情報について説明する。

【0035】特殊再生制御情報は、対象となる映像データの特殊再生のための制御情報であって、特殊再生制御情報生成装置により、当該映像データをもとにして作成され、当該映像データに付与される、あるいは関連付けられる制御情報である。特殊再生は、通常再生以外の方法による再生であり、例えば、倍速再生（あるいは高速再生）、飛び越し再生（あるいは飛び越し連続再生）、トリック再生等がある。トリック再生には、例えば、入れ替え再生、重複再生、スロー再生等、様々な種類がある。特殊再生制御情報は、例えば映像データを再生する映像再生装置において特殊再生を行う際に参照される。

【0036】図1に、特殊再生の対象となる映像データをもとにして作成される特殊再生制御情報の基本的なデータ構造の一例を示す。

【0037】このデータ構造は、複数のフレーム情報  $i$  ( $i = 1 \sim N$ ) を元映像データにおけるフレーム出現順序と対応付けて記述したもので、各フレーム情報は、特殊再生時に表示すべき元映像（1枚のフレーム、または複数枚の連続するフレーム群、もしくは複数枚の近接するフレーム群すなわち複数枚の連続するフレームのうち的一部分からなるフレーム群）の所在を示す情報を含む映像位置情報101と、その元映像についての特殊再生時に表示すべき表示時間を示す情報およびまたは表示時間を算出する基となる情報を含む表示時間制御情報102の組を含む。

【0038】図1は、フレーム情報  $i$  を映像データにおけるフレーム出現順序で配列して記述したものであるが、フレーム情報  $i$  内にそのフレーム情報の順序を示す情報を記述すれば、フレーム情報  $i$  をどのような順番で配列して記述しても構わない。

【0039】これらの複数のフレーム情報  $i$  からなるフレーム情報群に添付される再生倍率情報103は、特殊再生の再生速度の倍率を示し、フレーム情報に記述されている表示時間そのままではなく、フレームの表示時間を短縮し、数倍の速度で再生することを指定するために用いる。ただし、再生倍率情報103は必須の情報ではない。常に添付する構成と、常に添付しない構成と、添付するかどうかを個別的に選択可能とする構成とがある。再生倍率情報103が添付されていても特殊再生において必ずしもこれを用いなくても構わない。常に用い

る構成と、常に用いない構成と、用いるかどうかを個別的に選択可能とする構成とがある。

【0040】図1において、フレーム情報群に対して再生倍率情報と共に、または再生倍率情報に代えて更に他の制御情報を付加する構成も可能である。図1において、各フレーム情報  $i$  に更に他の制御情報を付加する構成も可能である。それらの場合において、映像再生装置側では特殊再生制御情報に含まれる各々の情報を全て使用してもよいし、それらの一部の情報のみを使用してもよい。

【0041】図2に、このような特殊再生制御情報の生成装置の構成例を示す。

【0042】図2に示されるように、この特殊再生制御情報生成装置は、映像データ記憶部2、映像位置情報処理部11および表示時間制御情報処理部12を含む映像データ処理部1、特殊再生制御情報記憶部3を備えている。詳しくは後述するが、図2の構成では、映像データを表示する場合は、元映像データ（符号化されている）をデコードして画像データとしてから表示するので、表示が指示されてから実際に画像が表示されるまでに、デコードのための処理時間がかかる。この時間を短縮するために、特殊再生に使う映像データを予めデコードしておき、画像データファイルを記憶しておく方法が考えられる。このような画像データファイルを用いる場合（常に画像データファイルを用いる形態をとる場合、または画像データファイルを用いるかどうか選択可能とする形態をとる場合）には、図3に示すように、映像データ処理部1内に画像データファイル作成部13を更に備え、映像データ処理部1に画像データファイル記憶部4を接続する。特殊再生制御情報に映像データをもとにして求めた他の制御情報を付加する場合には、適宜、映像データ処理部1内に該当する機能が追加される。

【0043】本処理においてユーザの操作を介入させる形態をとる場合には、例えば映像データをフレーム単位で表示させ、ユーザの指示入力等を受け付けるなどの機能を提供するGUIが用いられる（図2や図3では省略している）。

【0044】図2、図3においては、CPUやメモリ、必要に応じて設けられる外部記憶装置やネットワーク通信装置、必要に応じて使用されるドライバソフトやOS等のソフトウェアについては省略している。

【0045】映像データ記憶部2は、特殊再生制御情報、または特殊再生制御情報および画像データファイルを生成する処理対象となる映像データを記憶するためのものである。特殊再生制御情報記憶部3は、生成された特殊再生制御情報を記憶するためのものである。

【0046】画像データファイル記憶部4は、画像データファイル作成部13により作成された画像データファイルを記憶するためのものである。

【0047】映像データ記憶部2、特殊再生制御情報記

10

20

30

40

50

憶部3、画像データファイル記憶部4はいずれも、例えばハードディスクや光ディスクや半導体メモリなどで構成される。映像データ記憶部2、特殊再生制御情報記憶部3、画像データファイル記憶部4は、別々の記憶装置によって構成されていてもよいが、それらの全部または一部が同一の記憶装置によって構成されていてもよい。

【0048】映像データ処理部1は、処理対象となる映像データをもとにして、特殊再生制御情報（または特殊再生制御情報および画像データファイル）を生成するためのものである。

【0049】映像位置情報処理部11は、特殊再生時に表示すべきもしくは表示可能な映像フレーム（群）を決定（抽出）し、各フレーム情報iに記述すべき情報101を作成する処理を行う。

【0050】表示時間制御情報処理部102は、各フレーム情報に係る映像フレーム（群）の表示時間に関する情報102を作成する処理を行う。

【0051】画像データファイル作成部13は、映像データから各画像データファイルを作成する処理を行う。

【0052】特殊再生制御情報生成装置は、例えば、計算機上でソフトウェアを実行する形で実現することができる。特殊再生制御情報生成のための専用の装置として実現してもよい。

【0053】図4に、図2の構成の場合の制御情報生成手順の一例を示す。記憶部2から映像データを読み込み（ステップS11）、映像位置情報を作成し（ステップS12）、表示時間制御情報を作成し（ステップS13）、映像位置情報、表示時間制御情報からなる特殊再生制御情報を記憶部3に保存する（ステップS14）。図4の手順は、各フレーム情報ごとに逐次行ってもよいし、各処理をバッチ的に行ってもよい。その他の手順も可能である。

【0054】図5に、図3の構成の場合の制御情報生成手順の一例を示す。図4の手順に対して、さらに画像データファイルを作成し保存する手順が加わっている（ステップS22）。ここでは、画像データファイルの作成およびまたは保存は映像位置情報の作成とともに行っているが、図5とは異なるタイミングで行うことも可能である。図4の場合と同様、図5の手順は、各フレーム情報ごとに逐次行ってもよいし、各処理をバッチ的に行ってもよい。その他の手順も可能である。

【0055】次に、図6に、映像再生装置の構成例を示す。

【0056】図6に示されるように、この映像再生装置は、制御部21、通常再生処理部22、特殊再生処理部23、表示部24、コンテンツ記憶部25を備えている。映像データに音声などの音（Audio）が付加されているコンテンツを扱う場合には、音声出力部を備えていることが望ましい。映像データにテキストデータが付加されているコンテンツを扱う場合には、テキストは

表示部24に表示してもよいし、音声出力部から出力させることも可能である。プログラムが添付されているコンテンツを扱う場合には、添付プログラム実行部を設けるようにしても良い。

【0057】コンテンツ記憶部25には、少なくとも映像データと特殊再生制御情報が記憶されている。詳しくは後述するように、画像データファイルを用いる形態をとる場合には、更に画像データファイルが記憶されている。音声データやテキストデータや添付プログラムが更に記憶されていることもある。

【0058】コンテンツ記憶部25は、一箇所に集中配置されていても、複数箇所に分散配置されていてもよく、要は通常再生処理部22や特殊再生処理部23でアクセスすることができればよい。映像データや特殊再生制御情報や画像データファイルや音声データやテキストデータや添付プログラムは、別々の媒体に格納されていてもよいし、同一の媒体に格納されていてもよい。媒体としては、例えばDVDなどが用いられる。それらはネットワークを介して伝送されるデータであってもよい。

【0059】制御部21は、基本的には、GUI等のユーザ・インタフェースを介してユーザからコンテンツに対する通常再生や特殊再生などの指示を受け、該当する処理部分に、指定されたコンテンツについての指定された方法による再生を指示するなどの制御を行う。

【0060】通常再生処理部22は、指定されたコンテンツを通常再生するためのものである。

【0061】特殊再生処理部23は、指定されたコンテンツについて、特殊再生制御情報を参照して、指定された内容の特殊再生（例えば、倍速再生、飛び越し再生、トリック再生等）を行うためのものである。

【0062】表示部24は、映像を表示するためのものである。

【0063】映像再生装置は、例えば、計算機上でソフトウェアを実行する形で実現することができる（もちろん、一部分に、ハードウェアを用いてもよい（例えば、デコードボード（MPEG-2デコーダ）など）。映像再生のための専用の装置として実現してもよい。

【0064】図7に、図6の映像再生装置の再生処理手順の一例を示す。ステップS31でユーザから要求されたのは通常再生か特殊再生かを判断する。通常再生の要求があった場合は、ステップS32で指定された映像データを読み込み、ステップS33で通常再生を行う。ユーザから特殊再生の要求があった場合は、ステップS34で指定された映像データに対応する特殊再生制御情報を読み込み、ステップS35で表示すべき映像の位置の特定と、表示時間の決定を行い、ステップS36で該当するフレーム（群）を映像データ中（または画像データファイル）から読み込み、ステップS37で指定された内容の特殊再生を行う。表示すべき映像の位置の特定およびまたは表示時間の決定は、図7とは異なるタイミングで行う

ことも可能である。図7の特殊再生の手順は、各フレーム情報ごとに逐次行ってもよいし、各処理をバッチ的に行ってもよい。その他の手順も可能である。例えば、各フレームの表示時間を等しく一定の値にするような再生方法の場合には、表示時間の決定は行われない。

【0065】通常再生と特殊再生のいずれにおいても、ユーザが種々の指定（例えば、コンテンツにおける再生開始点およびまたはコンテンツにおける再生終了点、倍速再生における再生速度、倍速再生における再生時間、その他の特殊再生の方法、等）を要求できるようにするとより効果的である。

【0066】次に、特殊再生制御情報のフレーム情報の生成のアルゴリズムや、特殊再生時の表示時間の決定のアルゴリズムなどについて、概略的に説明する。

【0067】フレーム情報の生成時には、映像データのうちから特殊再生で使用するフレームの決定、映像位置情報の作成、表示時間制御情報の作成が行われる。

【0068】フレームの決定は、(1)当該映像データについての何らかの特徴量に基づいて行う方法（例えば、隣接フレーム間の特定の特徴量（例えば、フレーム間の画面変化量）が各抽出フレーム間でその総和が一定になるようにする方法、各抽出フレーム間での全フレームの重要度の総和が一定になるようにする方法）、

(2)画一的な基準により行う方法（例えば、ランダムに抽出する方法、等間隔に抽出する方法）、などがある。

【0069】表示時間制御情報の作成には、(i)表示時間または表示フレーム数の絶対値または相対値を求める方法、(ii)表示時間または表示フレーム数の基準となる情報（例えば、ユーザ指定、映像中の文字、映像に同期した音、映像中の人、あるいは映像中の特定パターン等）に基づいて得られる重要度）を求める方法、(iii)上記の(i)と(ii)の両方を記述する方法、などがある。

【0070】(1)または(2)と、(i)または(ii)または(iii)とは、適宜組み合わせることが可能である。もちろん、それ以外の方法も可能である。それらのうちの特定の1つの組み合わせのみ可能としてもよいし、それらのうちの複数の組み合わせを可能とし、適宜選択できるようにしてもよい。

【0071】特殊な場合として、(1)の方法でのフレームの決定と同時に(i)の表示時間または表示フレーム数の相対値が求まる方法がある。常にこの方法を用いる場合には、表示時間制御情報処理部102を省くことも可能である。

【0072】特殊再生時には、フレーム情報に含まれる(i)または(ii)または(iii)の表示時間制御情報を参照して行うことを想定しているが、記述されている値に従うようにしてもよいし、記述されている値を修正して使うようにしてもよいし、記述されている値またはこ

れを修正した値に加えて独自に用意した他の情報やユーザから入力された情報をも使うようにしてもよいし、独自に用意した他の情報やユーザから入力された情報のみをも使うようにしてもよい。それらのうちの複数の方法を可能とし、適宜選択できるようにしてもよい。

【0073】次に、特殊再生の概略について説明する。

【0074】倍速再生（あるいは高速再生）は、映像データ・コンテンツを構成する全フレームのうちの一部のフレームを再生することによって、もとのコンテンツを通常再生するのに要する時間より短い時間で再生を行うものである。例えば、フレーム情報で示されるフレームを、フレーム情報で示される表示時間ずつ、その時系列順に表示する。ユーザから、もとのコンテンツを通常再生する速度の何倍で再生するか（もとのコンテンツを通常再生するのに要する時間の何分の一の時間で再生するか）を指定する倍速指定や、どのくらいの時間をかけて再生するかを指定する時間指定などの要求を受け付け、該要求を満たすように各フレーム（群）の表示時間を求めて再生するようにしてもよい。そのため、この倍速再生は要約再生とも称する。

【0075】飛び越し再生（あるいは飛び越し連続再生）は、倍速再生において、例えば後述する再生／非再生情報に基づいて、フレーム情報で示されるフレームの一部を非再生とする。フレーム情報で示されるフレームのうち非再生とされたフレームを除いたフレームについて倍速再生するものである。

【0076】トリック再生は、通常再生以外の再生から、上記の倍速再生および飛び越し再生を除いたものである。例えば、フレーム情報で示されるフレームを再生する際に、ある部分について時系列順を入れ替えて再生する入れ替え再生、フレーム情報で示されるフレームを再生する際に、ある部分については複数回繰り返し再生する重複再生、フレーム情報で示されるフレームを再生する際に、ある部分については、他の部分より低速に再生し（通常再生時の速度にする場合と、通常再生時より低速にする場合とを含む）、あるいは他の部分より高速に再生し、あるいは一定時間表示して静止させ、あるいはそれらを適宜組み合わせる変速再生、フレーム情報で示されるフレームの一定の纏まりごとに時系列をランダムにして再生するランダム再生など、様々な形態のものが考えられる。

【0077】もちろん、複数種類の方法を適宜組み合わせたものも可能である。例えば、倍速再生時に、重要な部分については、複数回再生するとともに、再生速度を通常再生速度とする方法など、多彩なバリエーションが考えられる。

【0078】以下、本実施形態についてより具体的に詳しく説明する。

【0079】まず、フレームの決定のための画像データの特性値として隣接フレーム間の画面変化量を用いる場

合を例にとって説明する。

【0080】ここでは、1つのフレーム情報に、1つのフレームを対応させる場合について説明する。

【0081】図8に、対象となる映像データをもとにして作成される、特殊再生制御情報のデータ構造の一例を示す。

【0082】このデータ構造は、図1における表示時間制御情報102として（または表示時間制御情報102の代わりに）、絶対的なまたは相対的な表示時間を示す情報である表示時間情報121を記述するようにしたものである。表示時間制御情報102に重要度を記述する構成などについては後で説明する。

【0083】映像位置情報101は、当該映像の元映像フレームにおける位置を特定可能とする情報であり、フレーム番号（例えば先頭フレームからのシーケンス番号）やタイムスタンプなどのようにストリーム内の1フレームを特定できるものならどのようなものを用いても構わない。元映像ストリームから抜き出したフレームに対応する画像データを別ファイルとする場合は、そのファイル位置を特定する情報としてURLなどを用いてもよい。

【0084】表示時間情報121は、当該映像を表示する時間あるいはフレーム数を特定可能とする情報であり、実際に時間あるいはフレーム数を単位として記述する方法と、他のフレーム情報に記述されている表示時間情報との相対的な時間の長さの関係がわかるような相対値（例えば正規化された数値）を記述する方法とがある。後者の場合は、全体の総再生時間から、各映像の実際の再生時間を算出することになる。各映像について、表示の継続時間を記述するのではなく、特定のタイミングを起点とした（例えば最初の映像の開始時間を0とした）開始時間と終了時間の組み合わせでの記述や、開始時間と継続時間の組み合わせでの記述を用いてもよい。

【0085】特殊再生では、映像位置情報101により特定される位置に存在する映像を、表示時間情報121により特定される表示時間だけ再生することを、配列に含まれるフレーム情報の数だけ逐次行うことを基本とする。

【0086】開始時間と終了時間又は継続時間が指定されており、かつ、この指定に従う場合には、映像位置情報101により特定される位置に存在する映像を、表示時間情報121により特定される開始時間から終了時間まで再生することを、配列に含まれるフレーム情報の数だけ逐次行うことを基本とする。

【0087】再生倍率などのパラメータや、別の付加情報を用いることにより、記述された表示時間を加工して再生することも可能である。

【0088】次に、図9～図11を用いて、映像の位置情報の記述方法を説明する。

【0089】図9は、元映像フレームを参照する映像位

置情報の記述方法を説明する図である。

【0090】図9において、時間軸200は、特殊再生のためのフレーム情報を作成する対象となる元映像ストリームに対応し、画像201は映像ストリーム中の記述対象となる1フレームに対応する。時間軸202は、元映像ストリームから抜き出した画像201を使って特殊再生を行うときの映像の再生時間に対応し、表示時間203はその中に含まれる1つの画像201に対応する区間である。この場合には、例えば、画像201の位置を示す映像位置情報101と、表示時間203の長さを示す映像表示時間121との組がフレーム情報として記述される。前述のように、画像201の位置の記述はフレーム番号やタイムスタンプなど、元映像ストリーム内の1フレームを特定できるものならなんでもよい。このフレーム情報が他の画像についても同様に記述される。

【0091】図10は、画像データファイルを参照する映像位置情報の記述方法を説明する図である。

【0092】図9によって示される映像位置情報の記述方法は、特殊再生を行おうとする元映像データ内のフレームを直接参照するものであったが、図10によって示される映像位置情報の記述方法は、元映像ストリームから抜き出した単一フレーム302に対応する画像データ300を別のファイルに用意し、その位置を記述するものである。ファイル位置の記述方法は、例えば、URLなどを用いることにより、ローカルな記憶装置上に存在する場合でも、ネットワーク上に存在する場合でも同様に扱うことが可能である。この画像データファイルの位置を示す映像位置情報101と、対応する表示時間301の長さを示す映像表示時間121との組をフレーム情報として記述する。

【0093】元映像フレームとの対応が必要な場合は、記述したフレーム情報に対応する元映像の単一フレーム302を示す情報（例えば図9の場合における映像位置情報と同様のもの）をフレーム情報に含めればよい。この場合、フレーム情報は、映像位置情報、表示時間情報、元映像情報より構成されることとなる。もちろん、元映像情報は、必要がなければ記述する必要はない。

【0094】図10の方法によって記述される画像データの形態は、特に制約はないが、例えば、元映像のフレームをそのまま用いたり、縮小して用いたりするようにしてもよい。これは、元映像を展開する必要がないので、高速に再生処理を行うためにも有効である。

【0095】縮小画像の作成は、元映像ストリームがMPEG-1やMPEG-2などによって圧縮されている場合には、そのストリームを部分的に復号するだけで、高速に作成することができる。この手法は、フレーム内符号化されているIピクチャフレーム（フレーム内符号化フレーム）のDCT（離散コサイン変換）係数のみを復号し、その直流成分を用いることによって、縮小画像を作成する。

【0096】図10の記述方法では、画像データをそれぞれ別のファイルに格納していたが、これらのファイルはランダムアクセス可能な映像フォーマット（例えば、Motion JPEG）を持つ画像データ群格納ファイルにまとめて格納してもよい。この場合、画像データの位置は、画像データ群格納ファイルの位置を示すURLおよび画像データ群格納ファイル内での位置を示すフレーム番号またはタイムスタンプの組み合わせによって記述される。画像データ群格納ファイルの位置を示すURL情報は、個々のフレーム情報内に記述してもよいし、フレーム情報の配列外に付加情報として記述してもよい。

【0097】元映像のどのフレームを選択して画像データを作成して映像位置情報に記述するかについては、様々な方法をとることができる。例えば、元映像から等間隔に画像データを抽出してもよいし、画面の動きの多いところは狭い間隔で多くの画像データを抽出し、動きの少ないところは広い間隔で少ない画像データを抽出してもよい。

【0098】図11を参照しながら、フレーム選択方法の一例として、画面の動きに応じて、画面の動きの多いところは狭い間隔で多くの画像データを抽出し、動きの少ないところは広い間隔で少ない画像データを抽出する方法について説明する。

【0099】図11において、横軸はフレーム番号を表し、曲線800は（隣接フレーム間の）画面変化量の変化を表している。各フレームの画面変化量の算出方法は、後述する表示時間情報を求める際の手法と同様である。ここでは、画面の動きに応じて抽出間隔を決定するために、画像データ抽出元の映像フレーム間の画面変化量が一定となるような間隔を求める方法を示す。画像データ抽出元の映像フレーム間の画面変化量の合計を $S_i$ 、全フレームの画面変化量の総和を $S (= \sum S_i)$ とし、抽出する画像データ数を $n$ とする。画像データ抽出元フレーム間の画面変化量を一定にするには、 $S_i = S/n$ となればよい。図11では、画面変化量の曲線800が破線によって区切られた区間の面積 $S_i$ が一定になることに対応する。そこで、例えば、先頭フレームより、順次画面変化量を加算し、その値が $S/n$ を超えたフレームを画像データ抽出元の映像フレーム $F_i$ とする。

【0100】MPEGのIピクチャフレームより画像データを生成する場合には、算出された画像データ作成元フレームがIピクチャであるとは限らないので、近傍のIピクチャフレームより、画像データを作成する。

【0101】ところで、図11で説明した方法においては、画面変化量=0の区間に属する映像は、スキップされることになる。しかし、例えば静止画像が継続する場合には、重要な場面であることも多い。そこで、画面変化量=0が一定時間以上経過した場合には、そのときの

フレームを抜き出すようにしてもよい。この場合においては、例えば、先頭フレームより、順次画面変化量を加算し、その値が $S/n$ を超えたフレーム、または画面変化量=0が一定時間以上経過したフレームを画像データ抽出元の映像フレーム $F_i$ とするようにしてもよい。画面変化量=0が一定時間以上経過してフレームを抽出したときに、画面変化量の加算値を0にクリアする方法と、クリアせずに保持する方法とがある。この方法を使うか否かを選択可能にしてもよい。

【0102】図11の例の場合、いずれのフレームについても表示時間は同じとなるように表示時間情報121を記述することを想定しているが（この表示時間情報121に従って一定時間ずつ再生したときに、画面の変化量が一定となる）、表示時間情報121は一定ではなく、別の方法で求めて記述するようにしても構わない。

【0103】次に、1つのフレーム情報に、1または複数のフレームを対応させる場合について説明する。

【0104】この場合の特殊再生制御情報のデータ構造の一例は図8と同様である。

【0105】以下、図12～図20図21を用いて、映像の位置情報の記述方法を説明する。

【0106】図12は、元映像の連続するフレームを参照する映像位置情報の記述方法を説明する図である。

【0107】図9によって示される映像位置情報の記述方法は、特殊再生を行おうとする元映像内の1フレームを参照するものであったが、図12によって示される映像位置情報の記述方法は、元映像内の連続する複数のフレームの集合500を記述するものである。フレームの集合500は元映像内の連続する複数のフレームのうちの一部を抜き出したものであってもよい。また、フレームの集合500のうちに1つのフレームのみ含むものがあってもよい。

【0108】フレームの集合500が、元映像内の連続する複数のフレームまたは1つのフレームを含むものである場合には、フレーム位置の記述は、開始フレームおよび終了フレームの位置を記述するか、開始フレームの位置と記述区間の継続時間を記述する。1つのフレームを含むものである場合には、例えば、開始フレームと終了フレームの位置を同じにすればよい。位置や時間の記述は、フレーム番号やタイムスタンプなど、ストリーム内のフレームを特定できるものを用いる。

【0109】フレームの集合500が、元映像内の連続する複数のフレームのうちの一部である場合には、そのフレームが特定可能になるような情報を記述する。フレームの抜き出し方法が決まっており、例えば開始フレームおよび終了フレームの位置を記述すればフレームが特定可能となる場合には、それらを記述すればよい。

【0110】図12の表示時間情報501は、対応する元映像フレーム集合500に含まれるフレーム群全体に対応する総表示時間を示すものである。元映像フレーム

集合500に含まれる各フレームの表示時間については、特殊再生する装置側で適宜決定可能とすることができる。簡単な方法としては、上記の総表示時間を全フレーム数で均等割りして、1つのフレームの表示時間とする方法がある。もちろん、その他にも、種々の方法がある。

【0111】図13は、画像データファイルを参照する映像位置情報の記述方法を説明する図である。

【0112】図12によって示される映像位置情報の記述方法は、再生しようとする元映像内の連続するフレームを直接参照するものであったが、図13によって示される映像位置情報の記述方法は、元映像ストリームから抜き出したフレーム集合602に対応する画像データのフレーム集合600を別のファイルに用意し、その位置を記述するものである。ファイル位置の記述方法は、例えば、URLなどを用いることにより、ローカルな記憶装置上に存在する場合でも、ネットワーク上に存在する場合でも同様に扱うことが可能である。この画像データファイルの位置を示す映像位置情報101と、対応する表示時間601の長さを示す映像表示時間121との組をフレーム情報として記述する。

【0113】元映像フレームとの対応が必要な場合は、記述したフレーム情報に対応する元映像のフレーム集合602を示す情報（例えば図12の場合における映像位置情報と同様のもの）をフレーム情報に含めればよい。この場合、フレーム情報は、映像位置情報、表示時間情報、元映像情報より構成されることとなる。もちろん、元映像情報は、必要がなければ記述する必要はない。

【0114】画像データの形態や、画像データの作成、縮小画像の作成、画像データの格納方法、URLなどの位置情報の記述方法等については、前述と同様である。

【0115】元映像のどのフレームを選択して画像データを作成して映像位置情報に記述するかについても、前述と同様、様々な方法をとることができ、例えば、元映像から等間隔に画像データを抽出してもよいし、画面の動きの多いところは狭い間隔で多くの画像データを抽出し、動きの少ないところは広い間隔で少ない画像データを抽出してもよい。

【0116】上記した実施形態では画像データファイル300と元映像302の対応付けをフレーム単位で行っているが、元映像情報として記述するフレームの位置情報に時間的な幅を持たせることも可能である。そのため、フレーム情報のデータ構造は例えば、図14のようになる。図14では図8のフレーム情報に元映像情報3701が追加されている。元映像情報3701には特殊再生対象である元映像の対応区間の始点位置と区間長が、それぞれ始点情報3702、区間長情報3703として記述される。

【0117】元映像情報として記述する情報は区間を特定できるものならなんでもよい。ここでは始点位置と区

間長を用いたが、それらに代えて、始点位置と終点位置を元映像情報として用いてもよい。図15は図9に対して、元映像情報に時間的な幅を持たせた例である。この場合、例えば、同一のフレーム情報に含まれる映像位置情報、表示時間情報、元映像情報として、それぞれ、元映像フレーム3801の位置、表示時間3802、元映像フレーム区間3803（始点フレーム位置と区間長）を記述し、お互いが対応していることを示す。つまり、元映像フレーム区間3803を代表する画像として、映像位置情報に記述された元映像フレーム3801を表示することになる。

【0118】図16は図10に対して、元映像情報に時間的な幅を持たせた例である。この場合、例えば、同一のフレーム情報に含まれる映像位置情報、表示時間情報、元映像情報として、それぞれ、表示用画像データファイル3901の格納場所、表示時間3902、元映像フレーム区間3903（始点フレーム位置と区間長）を記述し、お互いが対応していることを示す。つまり、元映像フレーム区間3903を代表する画像として、映像位置情報に記述された画像データファイルの画像3901を表示することになる。

【0119】また、図12、13で示したようにフレームの集合を表示用映像として用いる場合において、表示用の映像に用いられている元映像フレーム区間とは異なる区間を元映像情報として対応付けても構わない。

【0120】図17は図12に対して、元映像情報に時間的な幅を持たせた例である。この場合、例えば、同一のフレーム情報に含まれる映像位置情報、表示時間情報、元映像情報として、それぞれ、元映像中のフレームの集合4001、表示時間4002、元映像フレーム区間4003（始点フレーム位置と区間長）を記述し、お互いが対応していることを示す。このとき、映像位置情報として記述するフレームの集合の区間4001と元映像情報として記述する元映像フレーム区間4003は必ずしも一致する必要はなく、異なる区間を表示用に用いても構わない。

【0121】図18は図13に対して、元映像情報に時間的な幅を持たせた例である。この場合、例えば、同一のフレーム情報に含まれる映像位置情報、表示時間情報、元映像情報として、それぞれ、表示に用いるフレーム集合4101の格納場所、表示時間4102、元映像フレーム区間4103（始点フレーム位置と区間長）を記述し、お互いが対応していることを示す。

【0122】このとき、映像位置情報として記述するフレームの集合4101の区間と元映像情報として記述する元映像フレーム区間4103は必ずしも一致する必要はない。つまり、表示用フレームの集合4101の区間が元映像フレーム区間4103より、短くてもよいし、長くてもよい。また、内容が全く異なる映像が含まれていてもよい。その他に、映像データファイルとして、元

映像情報に記述された区間のうち特に重要な区間のみを抜き出して、まとめた映像データファイルを使用する方法も考えられる。

【0123】これらのフレーム情報を用いて、例えば要約再生（特殊再生）映像を閲覧する際に、元映像中の対応するフレームを参照したい場合もある。

【0124】図19は要約表示された映像のフレームに対応する元映像のフレームから再生を開始するためのフローである。ステップS3601で、要約映像で再生開始フレームを指定する。ステップS3602では後述する方法で、指定されたフレームに対応する元映像フレームを算出する。ステップS3603では算出されたフレームより元映像を再生する。

【0125】もちろん、このフローは再生以外にも元映像の対応する位置を参照するために用いることが可能である。

【0126】ステップS3602において、対応する元映像フレームを算出する方法の一例として、要約映像で指定されたフレームの表示時間に対する比例配分を用いる方法を示す。i番目のフレーム情報の含まれる表示時間情報を $D_i$ 秒とし、元映像情報の区間始点位置を $t_i$ 秒、区間長を $d_i$ 秒とする。i番目のフレーム情報を用いた再生が始まってから、 $t$ 秒経過した位置を指定した場合、対応する元映像のフレームの位置は $T = t_i + d_i \times t / D_i$ となる。

【0127】次に、特殊再生、要約再生するフレーム選択の方法を説明する。

【0128】図20、図21を参照しながら、フレーム選択方法の一例として、画面の動きに応じて、画面の動きの多いところは狭い間隔で多くの画像データを抽出し、動きの少ないところは広い間隔で少ない画像データを抽出する方法について説明する。図20、図21の横軸や曲線800や $S_i$ や $F_i$ は図11と同様である。

【0129】図11の例では、画像データ抽出元フレーム間の画像変化量が一定となるような間隔で、1フレームずつ抽出した。図20、図21は、フレーム番号 $F_i$ を基準として複数のフレームの集合を抽出する例を示している。この場合、例えば、図20に示すようにフレーム番号 $F_i$ から一定数の連続するフレームを抽出するようにしてもよいし（フレーム長811とフレーム長812は同一）、図21に示すようにフレーム番号 $F_i$ から画像変化量の総和が一定となるようにそれぞれ該当する数の連続するフレームを抽出するようにしてもよい（面積813と面積814が同一）。もちろん、その他にも種々の方法が考えられる。

【0130】もちろん、前述した画面変化量=0が一定時間以上経過した場合の $F_i$ の抽出処理も用いることが可能である。

【0131】図11の場合と同様、図20、図21の例の場合、いずれのフレーム集合についても同じ表示時間

となるように表示時間情報121を記述するようにしてもよいし、別の方法で表示時間を求めて記述するようにしても構わない。

【0132】次に、表示時間を決定する処理の一例について説明する。

【0133】図22は、映像位置情報に記述された映像を、表示時間情報に記述された時間どおりに連続的に再生したときに、画面の変化量ができる限り一定となるような表示時間を求めるための基本処理手順の一例である。

【0134】この処理は、フレームの抽出をどのような方法で行った場合にも適用可能であるが、例えば図11のような方法でフレームを抽出した場合にはこの処理は省くことができる。何故ならば、図11は表示時間一定で画面の変化量ができる限り一定となるようにフレームを選択したからである。

【0135】ステップS71では、元映像の全フレームについて隣接フレームとの間の画面変化量を求める。映像の各フレームがビットマップにより表現されている場合は、隣接するフレーム間の画素の差分値を画面変化量とすることができる。映像がMPEGによって圧縮されている場合は、動きベクトルを用いて、画面変化量を求めることが可能である。

【0136】画面変化量の求め方の一例を説明する。

【0137】図23は、MPEGにより圧縮された映像ストリームから、全フレームの画面変化量を求めるための基本処理手順の一例である。

【0138】ステップS81では、Pピクチャのフレームから動きベクトルを抽出する。MPEGによって圧縮された映像のフレームは、図24に示すように、Iピクチャ（フレーム内符号化フレーム）、Pピクチャ（前方予測フレーム間符号化フレーム）、Bピクチャ（双方向予測フレーム間符号化フレーム）の並びによって記述される。このうち、Pピクチャには直前のIピクチャまたはPピクチャからの動きに対応する動きベクトルが含まれている。

【0139】ステップS82では、1つのPピクチャのフレームに含まれる各動きベクトルの大きさ（強度）を求め、その平均を直前のIピクチャまたはPピクチャからの画面変化量とする。

【0140】ステップS83では、Pピクチャのフレームに対して求めた画面変化量をもとに、Pピクチャ以外のフレームを含めた全フレームに対応する1フレームごとの画面変化量を算出する。例えば、Pピクチャのフレームの動きベクトルの平均値が $p$ で、参照元となる直前のIピクチャまたはPピクチャのフレームからの間隔が $d$ である場合、間の各フレームの1フレームあたりの画面変化量は $p/d$ である。

【0141】続いて、図22の手順におけるステップS72では、映像位置情報に記述する記述対象フレームか

ら、次の記述対象フレーム間での間にあるフレームの画面変化量の総和を求める。

【0142】図25は、1フレームごとの画面変化量の変化を記述した図である。横軸はフレーム番号に対応し、曲線1000が画面変化量の変化を表す。フレーム位置 $F_i$ の位置情報を持つ映像の表示時間を求める場合、次の記述対象フレーム位置である $F_{i+1}$ までの区間1001の画面変化量を累積加算する。これは、斜線部1002の面積 $S_i$ となり、フレーム位置 $F_i$ の動きの大きさと考えることができる。

【0143】続いて、図22の手順におけるステップS73では、各フレームの表示時間を求める。画面の変化量をできるだけ一定にするためには、画面の動きの大きいフレームほど、表示時間を多く配分すればよいので、各フレーム位置 $F_i$ の映像に配分する表示時間の再生時間に対する割合を、 $S_i / \sum S_i$ とすればよい。再生時間の総和を $T$ とすると、各映像の表示時間は、 $D_i = T \cdot S_i / \sum S_i$ となる。再生時間の総和 $T$ の値は、標準の再生時間で、元映像の総再生時間と規定しておく。

【0144】画面変化がなく $S_i = 0$ となる場合は、予め決められた下限値（例えば、1）を入れてもよいし、そのフレーム情報を記述しなくてもよい。 $S_i = 0$ とならないまでも、画面変化が非常に小さく、実際の再生においてほとんど表示されないことが予想されるフレームに関しても、下限値を代入してもよいし、フレーム情報を記述しなくてもよい。フレーム情報を記述しない場合は、 $S_i$ の値は $S_{i+1}$ に加算してもよいし、しなくてもよい。

【0145】この表示時間を求める処理は、特殊再生制御情報生成装置にてフレーム情報作成のために行うことができるが、映像再生装置側で特殊再生時に行うことも可能である。

【0146】次に、特殊再生を行う場合の処理の例について説明する。

【0147】図26は、記述された特殊再生制御情報に基づき、N倍速再生を行うための処理手順の一例である。

【0148】ステップS111では、再生倍率に基づいて、再生時の表示時間 $D'_i$ を算出する。フレーム情報に記述されている表示時間情報は、標準の表示時間なので、N倍速での再生を行う場合、各フレームの表示時間 $D'_i = D_i / N$ となる。

【0149】ステップS112では、表示のための初期化を行う。すなわち、先頭のフレーム情報を表示するように $i = 0$ とする。

【0150】ステップS113では、 $i$ 番目のフレーム情報の表示時間 $D'_i$ が予め設定された表示時間の閾値より大きいかなかを判定する。

【0151】大きい場合は、ステップS114において、 $i$ 番目のフレーム情報 $F_i$ に含まれる映像位置情報

の映像を $D'_i$ 秒間表示する。

【0152】大きくない（下回る）場合は、ステップS115に進み、表示時間の閾値を下回らない $i$ 番目のフレーム情報を順方向に探索する。この間、表示時間の閾値を下回ったフレーム情報の表示時間は、すべて探索の結果得られた $i$ 番目のフレーム情報の表示時間に加算し、表示時間の閾値を下回ったフレーム情報の表示時間は0とする。このような処理を行うのは、再生時の表示時間が非常に短くなると、表示する映像を準備する時間が表示時間よりも長くなり、表示が間に合わなくなる場合があるためである。そこで、表示時間が非常に短い場合は、表示をせずに先に進むようにする。その際に総再生時間が変わらないように、表示されなかった映像の表示時間を表示される映像の表示時間に加算する。

【0153】ステップS116では、まだ表示されていないフレーム情報が残っていないかを判断するために、 $i$ がフレーム情報の総数を下回っているか判定する。下回っている場合は、ステップS117へ進み、 $i$ を1増加させて次フレーム情報の表示を行う準備をする。 $i$ がフレーム情報の総数に到達した場合は、再生処理を終了する。

【0154】図27は、既定の表示サイクル（例えば、1秒間に30フレームを表示する場合、1表示サイクルは1/30秒）を基準にして、記述された特殊再生制御情報に基づき、N倍速再生を行うための処理手順の一例である。

【0155】ステップS121では、N倍速再生時に、各フレームの表示時間 $D'_i$ を、 $D'_i = D_i / N$ として求める。ここで算出される表示時間は、実際には表示サイクルとの関係があるので、算出された表示時間で映像を表示できるとは限らない。

【0156】図28は、算出された表示時間と表示サイクルの関係を表した図である。時間軸1300は算出された表示時間を示し、時間軸1301は表示レートに基づく表示サイクルを示す。表示レートが $f$ フレーム/秒の場合、表示サイクルの間隔は $1/f$ 秒となる。

【0157】したがって、ステップS122では、表示サイクルの開始点が含まれるフレーム情報 $F_i$ を探索し、ステップS123では、フレーム情報 $F_i$ に含まれる映像を1表示サイクル（ $1/f$ 秒）表示する。

【0158】例えば、表示サイクル1302は、表示開始点1303が、算出された表示時間1304に含まれるので、この表示時間に対応するフレーム情報の映像を表示する。

【0159】表示サイクルとフレーム情報との対応付け方法は、図29に示すように、表示サイクル開始点の最も近傍の映像を表示するようにしてもよい。図28の表示時間1305のように、表示時間が表示サイクルより小さくなった場合は、その映像の表示を省略してもよいし、強制的に表示してもよい。強制的に表示した場合

10

20

30

40

50



は、前後の表示時間を短くして全体の総表示時間が変わらないように調整する。

【0160】ステップS124では、現在の表示が最終表示サイクルであるかを調べ、最終表示サイクルであれば処理を終了し、最終表示サイクルでなければ次の表示サイクルを処理するために、ステップS125へ進む。

【0161】フレーム情報記述の他の例を説明する。

【0162】図8あるいは図14のデータ構造に含まれるフレーム情報は単一の元映像を要約する場合について扱ったものであるが、フレーム情報を拡張することによって、複数の元映像をまとめて要約することができる。図30はその一例で、個々のフレーム情報に含まれる元映像情報4201に元映像ファイルの位置などを示す元映像位置情報4202を追加した構造となっている。元映像位置情報に記述されるファイルは必ずしもファイル全体の区間を扱う必要はなく、一部区間のみを抜き出した形で用いてもよい。この場合、ファイル名などファイルの情報だけでなく、ファイルのどの区間が対象となっているかを示すための区間情報も合わせて記述する。映像ファイルから選択する区間は1つの映像に対して、複数であってもよい。

【0163】また、元映像が何種類か存在し、個々に識別情報が付与されている場合は、元映像位置情報の代わりに元映像識別情報を記述してもよい。

【0164】図31は元映像位置情報を追加したフレーム情報を用いて、複数の元映像をまとめて要約表示する例について説明する図である。この例では3つの映像（映像1、映像2、映像3）をまとめて、1つの要約映像を表示している。映像2に関しては全区間ではなく、4301と4302の2箇所の区間を取り出して、別々の元映像として扱っている。フレーム情報としてはこれらの元映像情報と共に、それぞれを代表する画像のフレーム位置（4301に対しては4303）が映像位置情報として、また、表示時間（4301に対しては4304）が表示時間情報として記述される。

【0165】図32は元映像位置情報を追加したフレーム情報を用いて、複数の元映像をまとめて要約表示する別の例について説明する図である。この例でも3つの映像をまとめて、1つの要約映像を表示している。映像2に関しては全区間ではなく、一部区間を取り出して、別々の元映像として扱っている。もちろん、図31のように複数の区間を取り出してもよい。フレーム情報としてはこれらの元映像情報（例えば映像2に加え4401の区間情報）と共に、それぞれを代表する画像ファイル（4402）の格納場所が映像位置情報として、また、表示時間（4403）が表示時間情報として記述される。

【0166】これらの例で説明したようなフレーム情報への元映像位置情報の追加は、フレームの集合を映像位置情報として用いる場合においても、全く同じように適

用することができ、複数の元映像をまとめた要約表示が可能である。

【0167】図33はフレーム情報を記述するための別のデータ構造である。このデータ構造では既に説明した映像位置情報、表示時間情報、元映像情報に加えて、動き情報4501と注目領域情報4502が加わっている。動き情報とはフレーム情報が対応する元映像の区間（すなわち元映像情報に記述された区間）の動きの大きさ（画面の変化量）を記述する。注目領域情報とは映像位置情報に記述されている画像の中で特に注目すべき領域の情報を記述したものである。

【0168】動き情報は図22において、映像の動きから表示時間を算出する際に用いたように映像位置情報に記述される画像の表示時間を算出するために用いることが可能である。この場合、表示時間情報を省略し、動き情報のみを記述しても、表示時間を記述した場合と同様に早送りなどの特殊再生を行うことができる（この場合、再生時に表示時間を計算する）。

【0169】表示時間情報と動き情報の両方を同時に記述することも可能であり、その場合は表示を行うアプリケーションが処理に合わせて必要な方を用いたり、組み合わせて用いられよい。

【0170】例えば、表示時間情報には動きと関係なく算出された表示時間を記述しておく。元映像から、重要な場面を切り出す表示時間の算出方法などがこれに該当する。このように算出された要約表示の早送りを行う際に、動き情報を用いて、動きの大きい部分は遅めに、動きの小さい部分の速めに再生を行うことによって、見落としの少ない早送りが可能である。

【0171】注目領域情報はフレーム情報の映像位置情報に記述された画像の中で特に注目すべき領域が存在するときに用いる。例えば、視聴者にとって重要と思われる人物の顔などがこれに該当する。このような注目領域情報を含む画像を表示する際には領域が分かるように矩形などを重ね合わせて表示してもよい。この表示は必須ではなく、そのまま画像を表示するだけでも構わない。

【0172】注目領域情報はフレーム情報などの特殊再生情報を加工して表示したりすることも可能である。例えば、一部のフレーム情報のみを再生表示する場合に、注目領域情報が含まれるフレーム情報を優先的に表示する。また、大きな面積をもつ矩形領域が含まれるほど、重要度が高いという解釈を用いて、選択表示することも可能である。

【0173】以上、画面変化量に基づいて要約再生するフレームを選択する場合を説明してきたが、以下では、重要度情報を利用してフレームを選択する場合を説明する。

【0174】図34は、映像に付帯させるフレーム情報のデータ構造の一例である。

【0175】このデータ構造は、図1のフレーム情報の

10

20

30

40

50

データ構造において、表示時間制御情報102として（または表示時間制御情報102の代わりに）、表示時間の基となる情報である重要度情報122を記述するようにしたものである。

【0176】重要度情報122は、対応するフレーム（またはフレーム集合）の重要度を表す。重要度は、例えば、一定範囲（例えば0から100の間）の整数として表現したり、一定範囲（例えば0から1の間）の実数として表現する。あるいは、上限を定めずに整数、実数値として表現しても良い。重要度情報は、映像の全てのフレームに対して付帯させても良いし、重要度の変化したフレームのみ付帯させても良い。

【0177】この場合、映像の位置情報の記述方法は、図9、図10、図12、図13のいずれの形態をとることも可能である。図11や図20、図21のフレーム抽出方法も利用可能である（この場合には、図11や図20、図21の画面変化量を重要度に置き換えればよい）。

【0178】次に、先に説明した例では、画面の変化量により表示時間の設定を行ったが、重要度情報により表示時間を設定することも可能である。以下、このような表示時間の設定方法について説明する。

【0179】先に例示した画面の変化量に基づく表示時間設定では、映像内容を理解しやすくするため、変化量の大きいところでは表示時間を長く設定し、変化量の小さいところでは表示時間を短く設定した。この重要度に基づく表示時間設定では、重要度の高いところは表示時間を長く設定し、重要度の低いところでは表示時間を短くすれば良い。すなわち、重要度による表示時間の設定方法は、基本的に画面の変化量に基づく表示時間設定方法（図25参照）と同様であるため、ここでは簡単に説明することにする。

【0180】図36に、この場合の基本処理手順の一例を示す。

【0181】ステップS191では、元映像の全フレームの重要度を求める。その具体的な方法については後で例示する。

【0182】ステップS192では、映像位置情報に記述する記述対象フレームから、次の記述対象フレームまでの間にあるフレームの重要度の総和を求める。

【0183】図37は、1フレームごとの重要度の変化を記述した図である。2200が重要度である。フレーム位置 $F_i$ の位置情報を持つ映像の表示時間を求める場合、次の記述対象フレーム位置である $F_{i+1}$ までの区間2201の重要度を加算する。加算結果は、斜線部2202の面積 $S'_i$ となる。

【0184】ステップS193では、各フレームの表示時間を求める。各フレーム位置 $F_i$ の映像に配分する表示時間の再生時間に対する割合を、 $S'_i / \sum S'_i$ とする。再生時間の総和を $T$ とすると、各映像の表示時

間は、 $D_i = T \cdot S'_i / \sum S'_i$ となる。再生時間の総和 $T$ の値は、標準の再生時間で、元映像の総再生時間と規定しておく。

【0185】重要度の和が $S'_i = 0$ となる場合は、予め決められた下限値（例えば、1）を入れてもよいし、そのフレーム情報を記述しなくてもよい。 $S'_i = 0$ とならないまでも、重要度が非常に小さく、実際の再生においてほとんど表示されないことが予想されるフレームに関しても、下限値を代入してもよいし、フレーム情報を記述しなくてもよい。フレーム情報を記述しない場合は、 $S'_i$ の値は $S'_{i+1}$ に加算してもよいし、しなくてもよい。

【0186】図35のように、図1のフレーム情報のデータ構造において、各フレーム情報 $i$ に、映像位置情報101と、表示時間情報121と、重要度情報122を記述するようにしてもよい。この場合において、特殊再生時には、表示時間情報121を用いるが重要度情報122を用いない方法と、重要度情報122を用いるが表示時間情報121を用いない方法と、両方用いる方法と、両方用いない方法がある。

【0187】この表示時間を求める処理は、特殊再生制御情報生成装置にてフレーム情報作成のために行うことができるが、映像再生装置側で特殊再生時に行うことも可能である。

【0188】次に、各フレームまたは場面（映像区間）の重要度の決定方法（例えば、図36のステップS191）について説明する。

【0189】映像のある場面が重要かどうかは、通常、様々な要因が絡み合っているため、重要度を決定する最も妥当な方法は、人間が決定する方法である。この方法では、映像のそれぞれの場面、または一定の時間間隔ごとに重要度評価者が重要度を評価し、重要度データへの入力を行う。ここで言う重要度データとは、フレーム番号または時刻と、そのときの重要度の値との対応表のことである。重要度の評価が主観的になってしまうことを避けるためには、複数の重要度評価者に同一の映像を評価してもらい、各場面または各映像区間ごとに平均値（またはメジアンなどでも良い）を算出して最終的な重要度を決定する。このような人手による重要度データ入力は、言葉では表現できないようなあいまいな印象や複数の要素を重要度に加味することが可能である。

【0190】人間が決定する手間を省くためには、重要であると思われる映像場面に出現しそうな事象を考え、このような事象を自動で評価して重要度に変換する処理を利用するのが好ましい。以下、重要度の自動生成の例をいくつか示す。

【0191】図38は、音声レベルの大きな場面が重要であるとして、重要度データを自動算出する際の処理手順の一例である（図38は機能ブロック図としても成立する）。

【0192】ステップS210の音声レベル算出処理では、映像に付随している音声データが入力されると、各時刻における音声レベルを算出する。音声レベルは瞬時に大きく変化するため、ステップS210の音声レベル算出処理では平滑化等の処理を行っても良い。

【0193】ステップS211の重要度算出処理では、音声レベル算出処理の結果出力される音声レベルを重要度に変換する処理を行う。例えば、あらかじめ定められている最低音声レベルを0、最高音声レベルを100として入力された音声レベルを0から100の値に線形に変換する。最低音声レベル以下の場合は0、最高音声レベル以上の場合は100とする。重要度算出処理の結果、各時刻における重要度が決定され、重要度データとして出力される。

【0194】図39は、他の重要度レベル自動決定方法の処理手順例である（図39は機能ブロック図としても成立する）。

【0195】図39の処理は、映像に付随する音声の中に、あらかじめ登録されている重要単語が多く出現する場面を重要であると判断するものである。

【0196】ステップS220の音声認識処理では、映像に付随する音声データが入力されると、音声認識処理により人が話した言葉（単語）をテキストデータに変換する。

【0197】重要単語辞書221には、重要な場面に登場しそうな単語が登録されている。登録されている単語の重要さの度合いが異なっている場合には、登録単語ごとに重みを付加しておく。

【0198】ステップS222の単語照合処理では、音声認識処理の出力であるテキストデータと重要単語辞書221に登録されている単語を照合し、重要な単語が話されたかどうかを判定する。

【0199】ステップS223の重要度算出処理では、単語照合処理の結果から映像の各場面や各時刻における重要度を算出する。この計算には、重要単語の出現数、重要単語の重みが使われ、例えば重要単語の出現した時刻の周辺（または出現した場面）の重要度を一定値もしくは重要単語の重みに比例する値だけ上昇させるといった処理を行う。重要度算出処理の結果、各時刻における重要度が決定され、重要度データとして出力される。

【0200】全ての単語の重みを同一とした場合には、重要単語辞書221は不要となる。これは、多くの単語が話された場面は重要な場面であると想定していることに相当する。このとき、ステップS222の単語照合処理では、単に音声認識処理から出力される単語の数をカウントする処理を行う。単語数ではなく、文字数をカウントするようにしても良い。

【0201】図40は、さらに他の重要度レベル自動決定方法の処理手順例である（図40は機能ブロック図としても成立する）。

【0202】図40の処理は、映像中に登場するテロップに、あらかじめ登録されている重要単語が多く出現する場面を重要であると判断するものである。

【0203】ステップS230のテロップ認識処理では、映像中の文字位置を特定し、文字位置の映像領域を2値化して文字認識を行う。認識された結果は、テキストデータとして出力される。

【0204】重要単語辞書231は、図39の重要単語辞書221と同様のものである。

10 【0205】ステップS232の単語照合処理では、図39の手順におけるステップS222と同様に、テロップ認識処理の出力であるテキストデータと重要単語辞書231に登録されている単語を照合し、重要な単語が登場したかどうかを判定する。

【0206】ステップS233の重要度算出処理では、図39の手順におけるステップS223と同様に、重要単語の出現数、重要単語の重みから各場面または各時刻における重要度を算出する。重要度算出処理の結果、各時刻における重要度が決定され、重要度データとして出力される。

20 【0207】全ての単語の重みを同一とした場合には、重要単語辞書231は不要となる。これは、テロップとして多くの単語が出現した場面は重要な場面であると想定していることに相当する。このとき、ステップS232の単語照合処理では、単にテロップ認識処理から出力される単語の数をカウントする処理を行う。単語数ではなく、文字数をカウントするようにしても良い。

30 【0208】図41は、さらに他の重要度レベル自動決定方法の処理手順例である（図41は機能ブロック図としても成立する）。

【0209】図41の処理は、映像中に登場するテロップの文字が大きいほど重要な場面であると判断するものである。

【0210】ステップS240のテロップ検出処理では、映像中の文字列の位置を特定する処理を行う。

【0211】ステップS241の文字サイズ算出処理では、文字列から個々の文字を切り出し、文字の大きさ（面積）の平均値または最大値を算出する。

40 【0212】ステップS242の重要度算出処理では、文字サイズ算出処理の出力である文字サイズに比例した重要度を算出する。算出された重要度が大きすぎたり小さすぎたりした場合には、しきい値処理により重要度をあらかじめ決められた範囲内に収める処理も行う。重要度算出処理の結果、各時刻における重要度が決定され、重要度データとして出力される。

【0213】図42は、さらに他の重要度レベル自動決定方法の処理手順例である（図42は機能ブロック図としても成立する）。

50 【0214】図42の処理は、映像中に人間の顔が登場する場面は重要であると判断するものである。

【0215】ステップS250の顔検出処理では、映像中にある人間の顔らしい領域を検出する処理を行う。処理の結果として、人間の顔と判断された領域の数（顔の数）が出力される。顔の大きさ（面積）の情報も同時に出力するようにしても良い。

【0216】ステップS251の重要度算出処理では、顔検出処理の出力である顔の数を定数倍して重要度を算出する。顔検出処理の出力が顔の大きさ情報を含む場合には、重要度は顔の大きさとともに増大するように計算される。例えば、顔の面積を定数倍して重要度を算出する。重要度算出処理の結果、各時刻における重要度が決定され、重要度データとして出力される。

【0217】図43は、さらに他の重要度レベル自動決定方法の処理手順例である（図43は機能ブロック図としても成立する）。

【0218】図43の処理は、あらかじめ登録されている画像と類似した映像が登場する場面は重要である判断するものである。

【0219】重要シーン辞書260には、重要と判断すべき画像が登録されている。画像は生データとして記録されていたり、データ圧縮された形式で記録されている。画像そのものではなく、画像の特徴量（色ヒストグラムや周波数など）を記録しておいても良い。

【0220】ステップS261の類似度／非類似度算出処理では、重要シーンに登録されている画像と入力された画像データとの類似度または非類似度を算出する。非類似度としては、2乗誤差の総和や絶対値差分の総和などが用いられる。重要シーン辞書260に画像データが記録されている場合には、対応する画素ごとの2乗誤差の総和や絶対値差分の総和などが非類似度として算出される。重要シーン辞書260に画像の色ヒストグラムが記録されている場合には、入力された画像データに対して同様の色ヒストグラムを算出し、ヒストグラム同士の2乗誤差の総和や絶対値差分の総和を算出して非類似度とする。

【0221】ステップS262の重要度算出処理では、類似度／非類似度算出処理の出力である類似度または非類似度から重要度を算出する。類似度が入力される場合には類似度が大きいほど大きな重要度となるように、非類似度が入力される場合には非類似度が大きいほど小さな重要度となるように重要度は計算される。重要度算出処理の結果、各時刻における重要度が決定され、重要度データとして出力される。

【0222】さらに他の重要度レベル自動決定方法として、瞬間視聴率の高い場面を重要とする方法がある。瞬間視聴率のデータは、視聴率調査の集計結果として得られるものであり、この瞬間視聴率を定数倍することで重要度が算出される。もちろん、その他にも種々の方法がある。

【0223】重要度の算出処理は、単独で用いてもよい

し、複数を同時に用いて重要度を算出するようにしてもよい。後者の場合には、例えば、いくつかの異なる方法で一つの映像の重要度を算出し、最終的な重要度は平均値または最大値として算出するようにしてもよい。

【0224】以上では画面変化量や重要度を例にとりて説明を行ったが、画面変化量およびまたは重要度とともに、あるいは画面変化量および重要度に代えて、その他の1または複数種類の情報を用いる（フレーム情報に記述する）ことも可能である。

10 【0225】次に、フレーム情報（図1参照）に、再生／非再生の制御のための情報を付加する場合について説明する。

【0226】映像データ中における、特定の場面あるいは部分（例えばハイライトシーン）のみを再生したり、特定の人物が登場している場面あるいは部分のみを再生したいなどというように、映像の一部のみを見たいという要求がある。

20 【0227】この要求を満たすため、フレーム情報に、再生するか非再生にするかを制御するための再生／非再生情報を付加するようにしてもよい。これにより、再生側では、この再生／非再生情報に基づいて、映像の一部のみを再生したり、逆に映像の一部のみを再生しなかったりすることができる。

【0228】図44、図45、図46に、再生／非再生情報を付加したデータ構造例を示す。

30 【0229】図44は、図8のデータ構造例において、再生／非再生情報123を付加したものである。もちろん、図45、図46は、図34、図35のデータ構造に再生／非再生情報123を付加したものである。図示していないが、図1のデータ構造例において、再生／非再生情報を付加してもよい。

【0230】再生／非再生情報123は、再生するか非再生にするかの2値情報を指定する方法と、再生レベルのような連続値を指定する方法がある。

【0231】後者の場合には、例えば、再生時に再生レベルがある閾値以上だったら再生し、そうでなければ非再生とする。閾値は、例えば、ユーザが直接的にまたは間接的に指定可能としてもよい。

40 【0232】再生／非再生情報123は、独立した情報とし保持してもよいが、再生か非再生かを選択的に指定する場合において、表示時間情報121により示される表示時間が特定の値（例えば、0あるいは-1など）のときに非再生であるとすることも可能である。あるいは、重要度情報122により示される重要度が特定の値（例えば、0あるいは-1など）のときに非再生であるとすることも可能である。この場合には、再生／非再生情報123は付加しなくてよい。

50 【0233】再生か非再生かをレベル値で指定する場合においても、表示時間情報121およびまたは重要度情報122（ただし、重要度をレベル値で表す場合）で代

用することも可能である。

【0234】再生／非再生情報123を独立した情報として保持する場合は、データ量がその分増えるが、再生側で非再生指定部分を再生しないようにしてダイジェストを見ることもできるし、非再生指定部分も再生して映像の全部を見ることも可能となる（再生／非再生情報123を独立した情報として保持しない場合は、非再生指定部分も再生して映像の全部を見るためには、例えば0として指定されている表示時間を適宜変更する必要がある）。

【0235】再生／非再生情報123は、人間が入力してもよいし、なんらかの条件より決定してもよい。例えば、映像の動き情報から動きが一定値以上大きいときは再生、そうでなければ非再生とすれば、動きの激しいところのみ再生できるし、色情報から肌色が一定値より大きい小さいかから決定すれば人物がいるところのみ再生できる。音の大小によって決定する手法、あらかじめ入力されている再生プログラム情報から決定する手法も考えられる。重要度をなんらかの手法で決定しておき、重要度情報から再生／非再生情報123を生成してもよい。再生／非再生情報を連続値としたときは、これらの情報を適当な関数で再生／非再生情報に変換することによって求ればよい。

【0236】図47は、再生／非再生情報123に基づいて、再生／非再生の制御を行って再生した例を示す。

【0237】図47において、元映像2151を、 $F_1 \sim F_n$ で表される映像フレーム位置情報または映像フレーム群位置情報2153と、 $D_1 \sim D_n$ で表される表示時間情報2154に基づいて再生するとする。このとき、再生／非再生情報2155は、表示時間情報2154に付加されるものとする。この例において、 $D_1, D_2, D_4, D_6$ の区間が再生となり、それ以外の区間が非再生となった場合には、再生映像2152としては、 $D_1, D_2, D_4, D_6$ の区間が連続的に再生される（それ以外は非再生となる）。

【0238】例えば、再生映像のフレーム $F_i$ において、再生／非再生情報123が再生を示すものであったときの表示時間を $D^+$ 、非再生であったときの表示時間を $D^-$ としたとき、元映像の再生部分の総時間が $T^+$ であるとする、 $\Sigma_i D^+ = T^+$ になる。通常は、 $D^+$ は、元映像と等倍速に表示時間を設定しておく。あらかじめ決め事として暗黙の固定された倍速としても良いし、何倍速に設定するかの情報を記述しても良い。N倍速再生したい場合は、再生部分の表示時間 $D^+$ を $1/N$ 倍する。例えば、決められた時間 $D^+$ で再生を行うようにするためには、各再生部分の表示時間 $D^+$ を $D^+ / \Sigma_i D^+$ 、 $D^+$ 倍に加工して表示すれば良い。

【0239】フレーム情報に基づいて各フレーム（またはフレーム群）の表示時間を決定する場合に、決定された表示時間を調整するようにしてもよい。

【0240】決定された表示時間を調整しない方法では、非再生の区間が発生したことを考慮せずに決定された表示時間をそのまま用いるので、非再生の区間にもともと0を越える表示時間が割り当てられていた場合には、その分だけ全体の表示時間が短くなる。

【0241】決定された表示時間を調整する方法では、例えば、非再生の区間にもともと0を越える表示時間が割り当てられている場合には、非再生の区間を再生したときと全体の表示時間が同じになるように、再生する各フレーム（またはフレーム群）の表示時間に一定数を乗じて、調整を行う。

【0242】ユーザが、調整するか否かを選択可能としてもよい。

【0243】ユーザがN倍速再生を指定した場合にも、決定された表示時間を調整せずにN倍速再生の処理を行ってもよいし、決定された表示時間を上記のようにして調整した後の表示時間を基礎としてN倍速再生の処理を行ってもよい（前者の方が表示時間が短くなる）。

【0244】ユーザが全体の表示時間を指定可能としてもよい。この場合にも、例えば、指定された全体の表示時間になるように、再生する各フレーム（またはフレーム群）の表示時間に一定数を乗じて、調整を行うようにしてもよい。

【0245】図48は、再生／非再生情報123に基づいて映像の一部のみを再生する処理手順の一例を示す。

【0246】ステップS162で該フレームのフレーム情報（映像位置情報及び表示時間情報）を読み出し、ステップS163で表示時間情報内の再生／非再生情報より該フレームを再生するか、非再生とするかを判断する。

【0247】判断結果が再生であれば、ステップS164で表示時間分だけ該フレームを表示する。そうでなければ、そのフレームは再生せず、次のフレームの処理に移る。

【0248】ステップS161で再生すべき映像が終了したかどうかを判別し、映像が終了したら、再生処理も終了する。

【0249】ところで、ステップS163で該フレームを再生するか非再生にするかを判断するときには、単純に再生／非再生情報が再生であれば再生し、非再生であれば再生しないという以外に、ユーザーの好みによって非再生部を再生するか再生しないかを決定したいことがある。このときは、映像の再生前にあらかじめ非再生部を再生するか再生しないかをユーザープロファイルなどから決定しておき、非再生部を再生するときは必ずステップS164でフレームの再生を行うようにする。

【0250】その他にも、再生／非再生情報が再生レベルとして連続値として保存されていたときは、ユーザープロファイルから再生と非再生を区別する閾値を求め、再生／非再生情報が閾値を超えているかどうかで再生す

るか非再生にするかを判断するようにしてもよい。ユーザプロファイルを使う以外にも、例えば、各フレームごとに設定された重要度から閾値を計算したり、ユーザからあらかじめ、またはリアルタイムに再生するか再生しないかの情報を受け取ってもよい。

【0251】このように、フレーム情報に、再生するか非再生にするかを制御するための再生／非再生情報123を付加することによって、映像の一部のみを再生することが可能となり、ハイライトシーンのみを再生したり、興味有る人物や物体が出ているシーンのみを再生したりすることが可能となる。

【0252】次に、フレーム情報（図1参照）に、表示される映像に関連した映像以外のメディア（例えばテキストや音）の位置情報と、それらを表示もしくは再生する時間を付加情報とする場合の記述方法について説明する。

【0253】図8では各フレーム情報100に映像位置情報101と表示時間情報102が含まれ、図34では各フレーム情報100に映像位置情報101と重要度情報103が含まれ、図35では各フレーム情報100に映像位置情報101と表示時間情報102と重要度情報103が含まれ、図44、図45、図46では、さらに再生／非再生情報123が含まれてる例を示したが、それらのいずれにおいても、さらに、0以上の音位置情報2703、音再生時間情報2704、0以上のテキスト情報2705、テキスト表示時間情報2706（ただし、いずれかは1以上とする）を付加するようにしてもよい。

【0254】図49は、図8のデータ構造例に、1組の音位置情報2703／音再生時間情報2704と、N組のテキスト情報2705／テキスト表示時間情報2706を付加した場合の例である。

【0255】音は、音位置情報2703に格納した位置から音再生時間情報2704に格納した時間だけ再生を行う。再生の対象は、最初から映像に付帯していた音情報でもかまわないし、バックグラウンドミュージックなどを作成してあらたに付加してもかまわない。

【0256】テキストは、テキスト情報2705に格納したテキスト情報をテキスト表示時間情報2706に格納した時間だけ表示する。1つの映像フレームに対して複数のテキスト情報を付加してもよい。

【0257】音の再生とテキストの表示を開始する時刻は、関連付けられた映像フレームが表示されるのと同時である。音の再生時間とテキストの表示時間も関連付けられた映像フレームの表示時間以内となる。複数の映像フレームにわたって、連続した音を再生する場合には、音の位置情報と再生時間を連続するように設定すればよい。

【0258】このような方法によって、要約音声や要約テキストなども可能になる。

【0259】図50に、フレーム情報とは別に音情報を記述する方法の一例を示す。これは、特殊再生を行う際に、表示されている映像フレームに関連する音声を再生するためのデータ構造例である。再生する音声の所在を示す位置情報2801と、音声の再生を開始する時刻2802と、再生を継続する時間2803の組を1つの音情報2800とし、この音情報の配列として記述される。

【0260】図51に、テキスト情報を記述するためのデータ構造の一例を示す。図50の音情報と同様な構造を持ち、表示するテキストの文字コード2901と、表示開始時刻2902と、表示時間2903の組を1つのテキスト情報2900とし、このテキスト情報の配列として記述される。2901に該当する情報として文字コードの代わりに、その文字コードを保存した場所、あるいはその文字を画像として保存した場所などを指す位置情報を用いてもよい。

【0261】上記の音情報やテキスト情報は、映像フレームの表示と同期をとり、表示されている映像フレームもしくはその映像フレームが存在する一定の映像区間に関連のある情報として表示される。図52に示すように、音情報やテキスト情報は、時間軸3001が示す時間の経過にしたがって、再生あるいは表示が開始される。まず、映像3002は、各映像フレームが記述された順序で、記述された表示時間ずつ表示されることによって再生される。3005、3006、3007はそれぞれ映像フレームを表しており、所定の表示時間が割り当てられている。音3003は、各音情報に記述された再生開始時刻になると再生され、同様に記述された再生時間を過ぎると再生を停止する。図52に示すように、同時に複数の音3008と3009が再生されてもよい。テキスト3004も音と同様に、各テキスト情報に記述された表示開始時刻になると表示され、記述された表示時間を過ぎると表示を停止する。同時に複数のテキスト3010と3011を表示してもよい。

【0262】音の再生開始時刻およびテキストの表示開始時刻は、映像フレームの表示を行う時刻と一致している必要はない。音の再生時間およびテキストの表示時間は、映像フレームの表示時間と一致している必要はない。これらは、自由に設定することができ、逆に音の再生時間やテキストの表示時間に応じて、映像フレームの表示時間を変更しても良い。

【0263】これらは人間が手動で設定することも可能である。

【0264】なお、図50、図51は映像のフレーム情報とは別に記述し、映像と同期して再生・表示する例であるが、映像とは別個に音、テキスト情報のみを記述して、音、テキストの要約再生、要約表示を行なってもよい。

【0265】人間が決定する手間を省くためには、重要

であると思われる映像場面に出現しそうな事象を考え、このような事象を自動的に設定する処理を利用するのが好ましい。以下、自動設定の例をいくつか示す。

【0266】図53は、ショットと呼ばれる画面の切り替わりから次の切り替わりまでの連続した映像区間を求め、そのショットに含まれる映像フレームの表示時間の総和を音声の再生時間とする処理手順の一例を示す(図53は機能ブロック図としても成立する)。

【0267】ステップS3101において、映像からショットを検出する。これには、「ゆう度比検定を用いた MPEGビットストリームからの動画像カット検出手法(信学論, Vol. J82-D-11, No. 3, pp. 361-370, 1999)」などの方法を用いる。

【0268】ステップS3102において、映像フレームの位置情報を参照して、それぞれの映像フレームがどのショットに属しているかを調べる。さらに、映像フレームの表示時間の総和を取ることによって、それぞれのショットの表示時間を求める。

【0269】例えば、音の位置情報はショットの始まりに対応した音声の位置とし、音の再生開始時刻はそれぞれのショットに属する最初の映像フレームの表示時刻にあわせ、音の再生時間はそのショットの表示時間に等しくすればよい。あるいは、音の再生時間に応じて、それぞれのショットに含まれる映像フレームの表示時間を修正してもよい。ここではショットを検出したが、(フレーム情報に重要度情報を記述するデータ構造をとる場合には)映像フレームに対する重要度を用いて、その重要度がしきい値以上の区間を求め、その区間に含まれる音を再生してもよい。

【0270】求められた再生時間が一定基準に満たない場合には、当該音声は再生しないようにしてもよい。

【0271】図54は、ショットもしくは重要度の高い映像区間に対応する音声データから、音声認識によって重要な単語を取り出し、その単語もしくは単語が含まれる音声もしくは複数の単語を組み合わせた音声を再生する処理手順の一例を示す(図54は機能ブロック図としても成立する)。

【0272】ステップS3201において、ショットを検出する。ショットのかわりに前記重要度の高い映像区間を求めてもよい。

【0273】ステップS3202において、得られた映像区間に対応する音声データの区間に対して、音声認識を行う。

【0274】ステップS3203において、認識結果の中から、重要な単語を含む音声もしくは重要単語部分の音声を求める。重要単語を選択するには、重要単語辞書3204を参照する。

【0275】ステップS3205において、再生用の音声を作成する。重要単語を含む連続した音声をそのまま

用いてもよいし、重要単語のみを抽出してもよい。重要単語を複数組み合わせた音声を作成してもよい。

【0276】ステップS3206において、作成した音声の再生時間に応じて、映像フレームの表示時間を修正する。ただし、音の再生時間が映像フレームの表示時間内となるように、選択した単語の数を減らし、音の再生時間を短くしてもよい。

【0277】図55に、テロップからテキスト情報を取得する手順の一例を示す(図55は機能ブロック図としても成立する)。

【0278】図55の処理は、テキスト情報は、映像中に表示されるテロップあるいは音声から取得するものである。

【0279】ステップS3301において、映像内で表示されるテロップを読み取る。これには、例えば文献「堀修: “テロップ領域のための映像からの文字部抽出法”, CVIM114-17, pp. 129-136 (1999)」に記述されている方法等により、元映像中のテロップを自動抽出するか、人間がテロップを読み取って手入力する方法がある。

【0280】ステップS3302において、読み取ったテロップ文字列から重要な単語を取り出す。重要単語の判定には、重要単語辞書3303を用いる。もちろん、読み取ったテロップ文字列をそのままテキスト情報としてもよい。抽出した単語を並べ、重要単語のみでその映像区間を表す文章を構成し、テキスト情報としてもよい。

【0281】図56に、音声からテキスト情報を取得する処理手順の一例を示す(図56は機能ブロック図としても成立する)。

【0282】ステップS3401の音声認識処理によって、音声を認識する。

【0283】ステップS3402において、認識した音声データから重要な単語を取り出す。重要単語の判定には、重要単語辞書3403を用いる。もちろん、認識した音声データをそのままテキスト情報としてもよい。抽出した単語を並べ、重要単語のみでその映像区間を表す文章を構成し、テキスト情報としてもよい。

【0284】図57に、ショットもしくは重要度の高い映像区間からテロップ認識によって、テキスト情報を取り出し、テキスト情報を作成する処理手順の一例を示す(図57は機能ブロック図としても成立する)。

【0285】ステップS3501において、映像からショットを検出する。ショットではなく、重要度の高い区間を求めてもよい。

【0286】ステップS3502において、その映像区間中表示されるテロップを認識する。

【0287】ステップS3503において、重要単語辞書3504を用いて、重要な単語を抽出する。

【0288】ステップS3505において、表示用のテ

10

20

30

40

50

キストを作成する。これには、重要単語を含むテロップ文字列を用いてもよいし、重要単語のみ、もしくは重要単語を複数用いた文字列をテキスト情報としてもよい。音声認識によってテキスト情報を得る場合には、ステップS3502のテロップ認識処理の部分を音声認識処理におきかえ、音声データを入力とすればよい。テキスト情報は、そのテキストがテロップとして表示された映像フレーム、あるいは音声として再生された時刻の映像フレームに合わせて表示する。あるいは、その映像区間中のテキスト情報を一度に表示してもよい。

【0289】図58にテキスト情報の表示例を示す。図58の(a)のように、テキスト情報表示部3601と映像表示部3602に分けてもよいし、図58の(b)のように、テキスト情報を映像表示部3603に重ねて表示してもよい。

【0290】映像フレーム、音情報、テキスト情報それぞれの表示時間(再生時間)は、すべてのメディア情報が同期するように調整する。例えば、映像を倍速再生する際には、まず、前述の方法で重要な音声を抽出し、通常再生の2分の1の時間の音声情報を取得しておく。次に、それぞれの音声に関連した映像フレームに表示時間を割り当てる。画面の変化量が一定となるように映像フレームの表示時間を決めた場合には、音声の再生時間やテキストの表示時間は、それぞれ関連する映像フレームの表示時間内とする。もしくは、ショットのように複数の映像フレームを含む区間を求めておき、その区間に含まれる音声あるいはテキストをその区間の表示時間に応じて再生もしくは表示するようにする。

【0291】これまででは、映像データを中心に説明してきたが、音声データを中心に扱ったシステムももちろん可能である。

【0292】これまででは映像を中心とした要約表示について扱ってきたが、フレーム情報を持たない(すなわち映像を持たない)形式で、音情報やテキスト情報を用いることも可能である。この場合、元映像に対して、音情報とテキスト情報だけで構成される要約を作成することになる。また、音声データや音楽データに対して、音情報とテキスト情報だけで構成される要約を作成することも可能である。

【0293】その際に、フレーム情報の場合と同様に、音情報やテキスト情報に元の音声、音楽データとの対応関係を記述するための元データ情報を追加してもよい。

【0294】図59は図50に示すデータ構造の音情報に、元データ情報4901を含めたデータ構造の例である。元データ情報4901は入力映像の場合は映像の区間を示す時間(始点情報4902と区間長情報4903)となるし、入力音声/音楽の場合は音声/音楽の区間を示す時間となる。

【0295】図60は図30に相当するデータ構造の音情報に、元データ情報4901を含めたデータ構造の例

である。

【0296】図61は音情報を用いて、音声/音楽を要約した一例について説明したものである。この例では元となる音声/音楽をいくつかの区間に分割し、それぞれの区間の一部をその区間の要約音声/音楽として切り出して、要約を作成している。例えば、区間2の5001の部分を要約音声/音楽として切り出し、要約の5002の区間として再生する。区間を分割方法の例としては、音楽を楽章ごとに分けたり、会話を内容ごとに分けたりする方法が考えられる。

【0297】また、フレーム情報の場合と同様に、音情報やテキスト情報に元データファイルおよび区間の記述を含めることによって、複数の音声、音楽データをまとめて要約することもできる。このとき、個々の元データに識別情報が付与されている場合は元データファイルや区間を記述する代わりに、元データ識別情報を用いてもよい。

【0298】図62は音情報を用いて、音声/音楽を要約した一例について説明したものである。この例では複数の音声/音楽データに対して、それぞれ一部区間の要約音声/音楽として切り出して、要約を作成している。例えば、音声/音楽2の5101の部分を要約音声/音楽として切り出し、要約の5102の区間として再生する。1枚の音楽アルバムに含まれる曲の一部区間ずつを切り出してまとめ、試聴用の要約データを作成する用途などが考えられる。

【0299】アルバムなどを要約する場合など、曲名が分かった方がよい場合は音情報に音楽データの曲名を含めるようにしてもよい。もちろん、この情報は必須ではない。

【0300】次に、映像データや物体領域データの提供方法について説明する。

【0301】本実施形態の処理により作成された特殊再生制御情報がユーザの用に供される場合には、作成者側からユーザ側に何らかの方法で特殊再生制御情報を提供が必要がある。この提供の方法としても以下に例示するように種々の形態が考えられる。

(1) 映像データとその特殊再生制御情報とを1つ(または複数の)記録媒体に記録して同時に提供する形態

(2) 映像データを1つ(または複数の)記録媒体に記録して提供し、別途、特殊再生制御情報を1つ(または複数の)記録媒体に記録して提供する形態

(3) 映像データとその特殊再生制御情報とを同じ機会に通信媒体を介して提供する形態

(4) 映像データとその特殊再生制御情報とを異なる機会に通信媒体を介して提供する形態

これにより、映像コンテンツの特殊再生に供するための制御情報として、元映像から選択的に抽出したフレーム(群)の取得方法と、そのフレーム(群)に割り当てた表示時間の情報又はこれを得る基となる情報とを含むフ



レーム情報を、複数配列させて記述することにより、再生側では該制御情報に基づいた効果的な特殊再生が可能になる。

【0302】以上説明したように、本実施形態によれば、映像コンテンツに対する特殊再生に供するための特殊再生制御情報を記述する特殊再生制御情報記述方法において、前記映像コンテンツを構成する映像データの全フレーム系列のなかから選択的に抽出された1フレーム又は連続若しくは近接する複数フレームからなるフレーム群ごとに、該1フレーム又は該フレーム群のデータが存在する位置を示す第1の情報と、該1フレーム又は該フレーム群に対して付与された表示時間に関する第2の情報及び又は該フレーム情報に対応する前記1フレーム又は前記フレーム群に対して付与された重要度を示す第3の情報をフレーム情報として記述する。

【0303】また、本実施形態によれば、映像コンテンツを構成する映像データの全フレーム系列のなかから選択的に抽出された1フレーム又は連続若しくは近接する複数フレームからなるフレーム群ごとに記述された、該1フレーム又は該フレーム群のデータが存在する位置を示す第1の情報と、該1フレーム又は該フレーム群に対して付与された表示時間に関する第2の情報及び又は該1フレーム又は該フレーム群に対して付与された重要度を示す第3の情報とを含むフレーム情報を少なくとも含む特殊再生制御情報を格納したコンピュータ読取り可能な記録媒体も提供される。

【0304】さらに、映像コンテンツに対する特殊再生に供するための特殊再生制御情報を生成する特殊再生制御情報装置／生成方法の実施形態において、前記映像コンテンツを構成する映像データの全フレーム系列のなかから、特殊再生に供される一部のフレームを、1フレーム又は連続若しくは近接する複数フレームからなるフレーム群ごとに、該フレーム系列に沿って順次選択的に抽出し、抽出された前記1フレーム又は前記フレーム群ごとに、該1フレーム又は該フレーム群のデータが存在する位置を示す映像位置情報と、該1フレーム又は該フレーム群に対して割り当てるべき表示時間の情報又はこれを算出する基となる情報を含む表示時間制御情報とを生成し、前記1フレーム又は前記フレーム群ごとに生成された、前記映像位置情報及び前記表示時間制御情報をフレーム情報として記述することによって、前記映像コンテンツに対する特殊再生制御情報を作成する。

【0305】また、映像コンテンツに対する特殊再生を行うことが可能な映像再生装置／方法の実施形態において、前記映像コンテンツに付随する、該映像コンテンツを構成する映像データの全フレーム系列のなかから選択的に抽出された1フレーム又は連続若しくは近接する複数フレームからなるフレーム群ごとに記述された、該1フレーム又は該フレーム群のデータが存在する位置を示す映像位置情報と、該1フレーム又は該フレーム群に対

して割り当てるべき表示時間の情報又はこれを算出する基となる情報を示す表示時間制御情報とを含むフレーム情報を少なくとも含む特殊再生制御情報を参照し、前記フレーム情報に含まれる映像位置情報に基づいて、各フレーム情報に対応する前記1フレーム又は前記フレーム群のデータを取得するとともに、少なくとも各々の前記フレーム情報に含まれる前記表示時間制御情報に基づいて、各フレーム情報に対して割り当てるべき表示時間を決定し、取得された前記1フレーム又は複数フレームのデータを、決定された前記表示時間によって再生することを、所定の順序で行うことによって、特殊再生を行う。

【0306】本発明の実施形態では、例えば、あらかじめ表示に用いる有効な映像フレームの位置情報または元映像からフレーム単位で取り出した画像データを準備し、その映像フレーム位置情報または画像データの表示時間に関する情報を元映像とは別に準備する。元映像から取り出した映像フレームもしくは画像データを表示情報に基づいて、それらを連続表示することにより、倍速再生、トリック再生、飛び越し連続再生などの特殊再生を行うことができる。

【0307】例えば、高速に内容を確認するための倍速再生においては、表示画面の画面の変化ができるだけ一定になるように、動きの大きいところは表示時間を長く、動きの小さいところは表示時間を短くするように、あらかじめ表示時間を決めておくようにしてもよい。または、表示に用いる映像フレームまたは画像データの動きの多い部分からは多く、動きの少ないところは少なくするように、位置情報を決めても、同じ効果が得られる。全体としてユーザによって指定される倍速値または再生時間になるようにコントロールされる値を準備するようにしてもよい。長い映像も短い時間で、見やすい倍速再生等でみることができ、内容を短い時間で把握することができる。

【0308】例えば、表示時間を重要度に応じて、重要な場所は表示時間を長く、低い場所は短くすることにより、重要な場所を見落とすにくい再生も可能である。

【0309】例えば、全映像フレームを表示せず、部分的に映像の一部を省略することにより、重要な部分だけを効率良く再生するようにしてもよい。

【0310】本発明の実施形態によれば、映像コンテンツの特殊再生に供するための制御情報として、元映像から選択的に抽出したフレーム（群）の取得方法と、そのフレーム（群）に割り当てた（絶対的若しくは相対的な）表示時間の情報又はこれを得る基となる情報とを含むフレーム情報を、複数配列させて記述することにより、再生側では該制御情報に基づいた効果的な特殊再生が可能になる。

【0311】例えば、以上の各機能は、ソフトウェアとしても実現可能である。上記実施形態は、コンピュータ

10

20

30

40

50

に所定の手段を実行させるための、あるいはコンピュータを所定の手段として機能させるための、あるいはコンピュータに所定の機能を実現させるためのプログラムを記録したコンピュータ読取り可能な記録媒体としても実施することもできる。

【0312】各実施形態で例示した構成は一例であって、それ以外の構成を排除する趣旨のものではなく、例示した構成の一部を他のもので置き換えたり、例示した構成の一部を省いたり、例示した構成に別の機能を付加したり、それらを組み合わせたりすることなどによって得られる別の構成も可能である。例示した構成と論理的に等価な別の構成、例示した構成と論理的に等価な部分を含む別の構成、例示した構成の要部と論理的に等価な別の構成なども可能である。例示した構成と同一もしくは類似の目的を達成する別の構成、例示した構成と同一もしくは類似の効果を奏する別の構成なども可能である。各実施形態内において、各種構成部分についての各種バリエーションは、適宜組み合わせることで実施することが可能である。各実施形態は適宜組み合わせることで実施することが可能である。各実施形態は、情報の記述方法としての発明、記述された情報としての発明、装置またはそれに対応する方法としての発明、装置内部またはそれに対応する方法としての発明等、種々の観点、段階、概念またはカテゴリに係る発明を包含・内在するものである。また、本発明は、コンピュータに所定の手段を実行させるための、あるいはコンピュータを所定の手段として機能させるための、あるいはコンピュータに所定の機能を実現させるためのプログラムを記録したコンピュータ読取り可能な記録媒体としても実施することもできる。

【0313】従って、この発明の実施の形態に開示した内容からは、例示した構成に限定されことなく発明を抽出することができるものである。

【0314】

【発明の効果】以上説明したように本発明によれば、利用者にとってより効果的な特殊再生を可能とするフレーム情報記述方法、フレーム情報生成装置及び方法並びに映像再生装置及び方法を提供することができる。

【図面の簡単な説明】

【図 1】本発明の一実施形態に係る特殊再生制御情報のデータ構造例を示す図。

【図 2】特殊再生制御情報生成装置の構成例を示す図。

【図 3】特殊再生制御情報生成装置の他の構成例を示す図。

【図 4】図 2 の構成の場合の処理手順の一例を示すフローチャート。

【図 5】図 3 の構成の場合の処理手順の一例を示すフローチャート。

【図 6】映像再生装置の構成例を示す図。

【図 7】図 6 の構成の場合の処理手順の一例を示すフローチャート。

【図 8】特殊再生制御情報のデータ構造例を示す図。

【図 9】元映像フレームを参照する映像位置情報について説明する図。

【図 10】画像データファイルを参照する映像位置情報について説明する図。

【図 11】画面の動きに応じた画像データの抽出方法について説明する図。

【図 12】元映像フレームを参照する映像位置情報について説明する図。

10 【図 13】画像データファイルを参照する映像位置情報について説明する図。

【図 14】元映像情報として記述するフレームの位置情報に時間的な幅を持たせる場合のフレーム情報のデータ構造を示す図。

【図 15】元映像フレームを参照する映像位置情報に時間的な幅を持たせた例を説明する図。

【図 16】画像データファイルを参照する映像位置情報に時間的な幅を持たせた例を説明する図。

20 【図 17】元映像フレームを参照する映像位置情報に時間的な幅を持たせた例を説明する図。

【図 18】元映像フレームを参照する画像データファイルに時間的な幅を持たせた例を説明する図。

【図 19】要約表示された映像のフレームに対応する元映像のフレームから再生を開始するためのフロー図。

【図 20】画面の動きに応じた画像データの抽出方法について説明する図。

【図 21】画面の動きに応じた画像データの抽出方法について説明する図。

30 【図 22】画面の変化量ができるだけ一定となる表示時間を求める処理手順の一例を示すフローチャート。

【図 23】MPEG 映像から全フレームの画面変化量を求める処理手順の一例を示すフローチャート。

【図 24】MPEG ストリームからの画像変化量の算出方法について説明する図。

【図 25】画面の変化量ができるだけ一定となる表示時間を求める処理手法について説明する図。

【図 26】特殊再生制御情報に基づく特殊再生を行う処理手順の一例を示すフローチャート。

40 【図 27】表示サイクルを基準に特殊再生を行う処理手順の一例を示すフローチャート。

【図 28】算出された表示時間と表示サイクルの関係について説明する図。

【図 29】算出された表示時間と表示サイクルの関係について説明する図。

【図 30】元映像位置情報を持つ特殊再生制御情報のデータ構造例を示す図。

【図 31】複数の元映像をまとめて要約表示する際の元映像フレームを参照する映像位置情報を説明する図。

50 【図 32】複数の元映像をまとめて要約表示する際の画像データファイルを参照する映像位置情報を説明する

図。

【図33】フレーム情報を記述するための別のデータ構造を示す図。

【図34】特殊再生制御情報のデータ構造例を示す図。

【図35】特殊再生制御情報のデータ構造例を示す図。

【図36】重要度から表示時間を求める処理手順の一例を示すフローチャート。

【図37】重要度から表示時間を求める手法について説明する図。

【図38】音声レベルの大きな場面を重要として重要度データを算出する処理手順の一例を示すフローチャート。

【図39】音声認識により重要な単語が多く出現している場面を重要として重要度データを算出する処理または時間あたりに話された単語の数が多い場面を重要として重要度データを算出する処理の手順の一例を示すフローチャート。

【図40】テロップ認識により重要な単語が多く出現している場面を重要として重要度データを算出する処理または時間あたりに出現したテロップに含まれる単語の数が多い場面を重要として重要度データを算出する処理の手順の一例を示すフローチャート。

【図41】大きな文字がテロップとして出現した場面を重要として重要度データを算出する処理の手順の一例を示すフローチャート。

【図42】人間の顔が多く登場する場面を重要として重要度データを算出する処理または人間の顔が大きく写る場面を重要として重要度データを算出する処理の手順の一例を示すフローチャート。

【図43】登録しておいた重要シーンと類似した映像が出現する場面を重要として重要度データを算出する処理の手順の一例を示すフローチャート。

【図44】特殊再生制御情報のデータ構造例を示す図。

【図45】特殊再生制御情報のデータ構造例を示す図。

【図46】特殊再生制御情報のデータ構造例を示す図。

【図47】再生するか非再生にするかの情報と再生映像の関係について説明する図。

【図48】再生／非再生判断を含む特殊再生の処理の手順の一例を示すフローチャート。

\*

\*【図49】音情報、テキスト情報を付加したときのデータ構造の一例を示す図。

【図50】音情報のみをフレーム情報とは別に記述するためのデータ構造の一例を示す図。

【図51】テキスト情報のみをフレーム情報とは別に記述するためのデータ構造の一例を示す図。

【図52】各メディア再生の同期について説明する図。

【図53】映像区間内での音再生開始時刻と音再生時間の決定手順の一例を示すフローチャート。

【図54】再生用音声データの作成と映像フレーム表示時間の修正の処理の手順の一例を示すフローチャート。

【図55】テロップ認識によるテキスト情報取得処理の手順の一例を示すフローチャート。

【図56】音声認識によるテキスト情報の取得処理の手順の一例を示すフローチャート。

【図57】テキスト情報の作成処理の手順の一例を示すフローチャート。

【図58】テキスト情報の表示方法について説明する図。

【図59】フレーム情報を記述するための別のデータ構造を示す図。

【図60】フレーム情報を記述するための別のデータ構造を示す図。

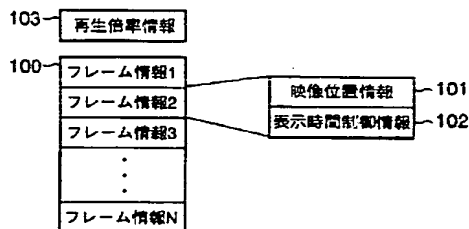
【図61】音楽データの要約再生を説明する図。

【図62】複数の音楽データの要約再生を説明する図。

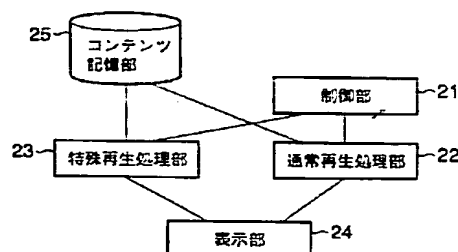
【符号の説明】

- 1…映像データ処理部
- 2…映像データ記憶部
- 3…特殊再生制御情報記憶部
- 4…画像データファイル記憶部
- 11…映像位置情報処理部
- 12…表示時間制御情報処理部
- 13…画像データファイル作成部
- 21…制御部
- 22…通常再生処理部
- 23…特殊再生処理部
- 24…表示部
- 25…コンテンツ記憶部

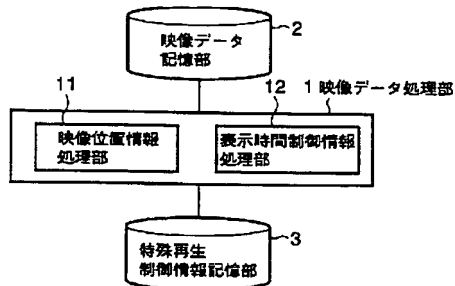
【図1】



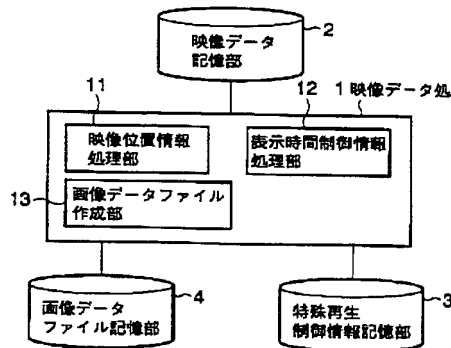
【図6】



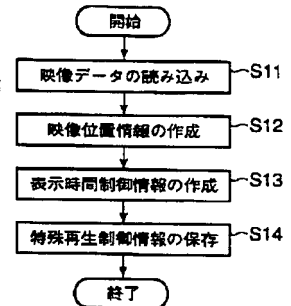
【図2】



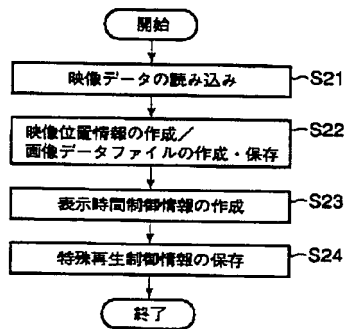
【図3】



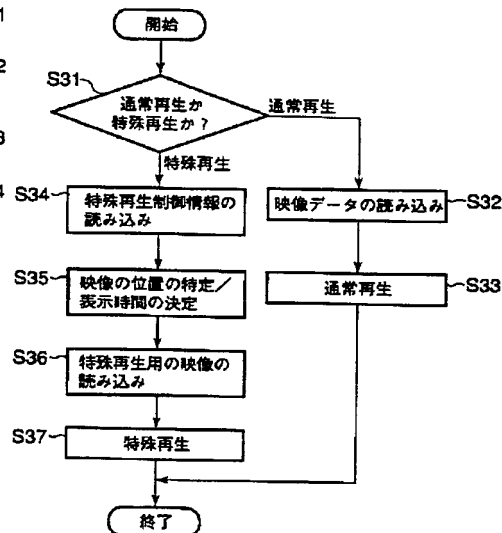
【図4】



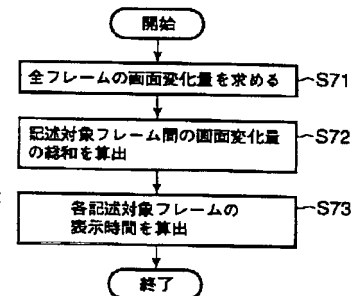
【図5】



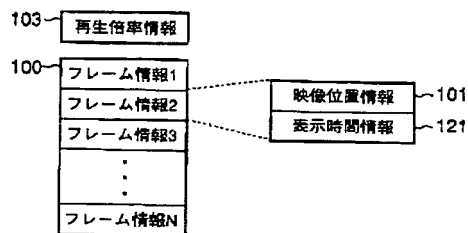
【図7】



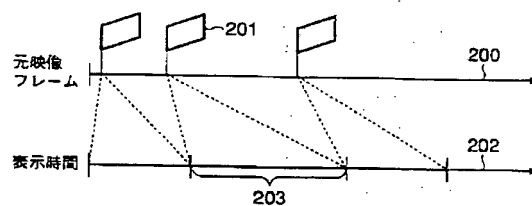
【図22】



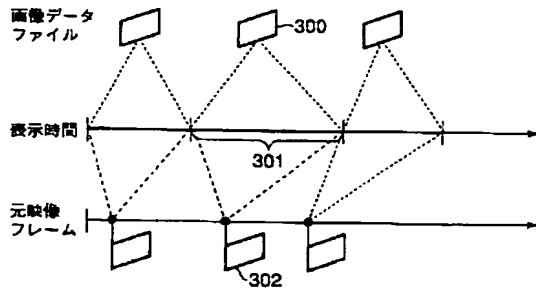
【図8】



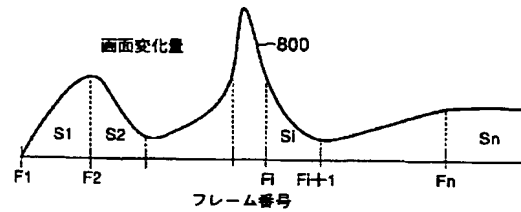
【図9】



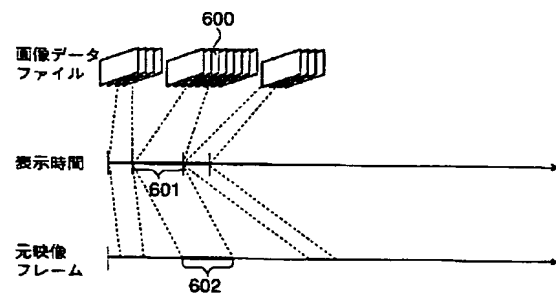
【図10】



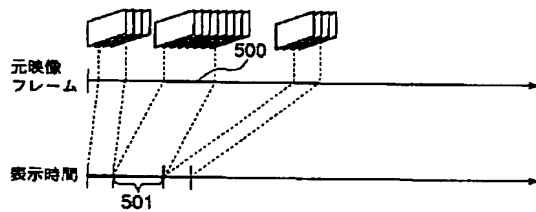
【図11】



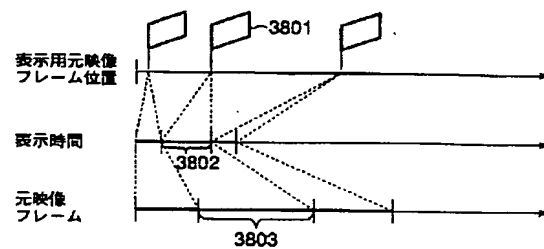
【図13】



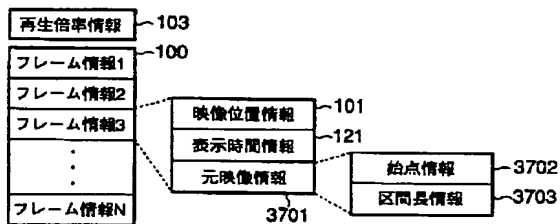
【図12】



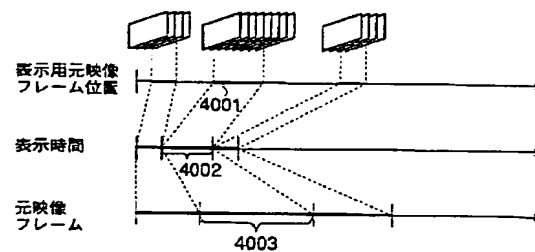
【図15】



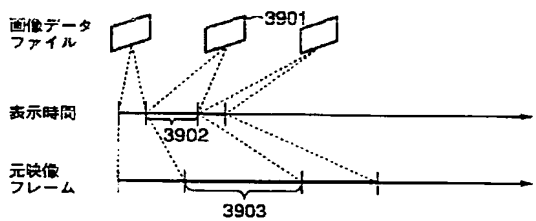
【図14】



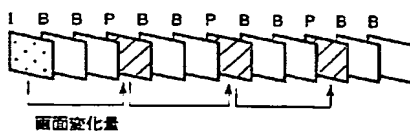
【図17】



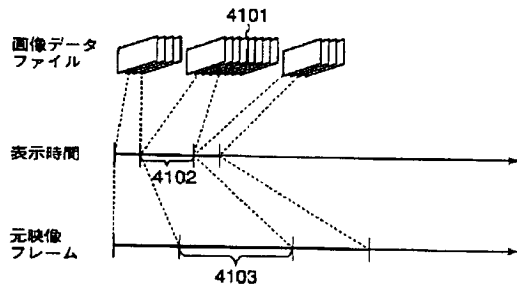
【図16】



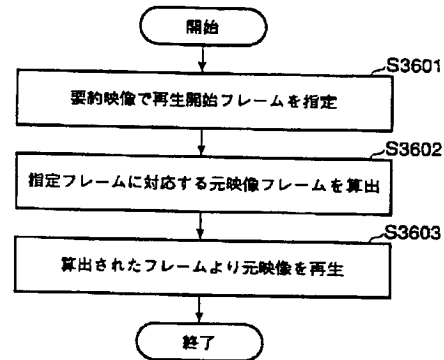
【図24】



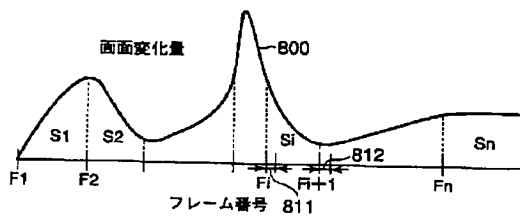
【図18】



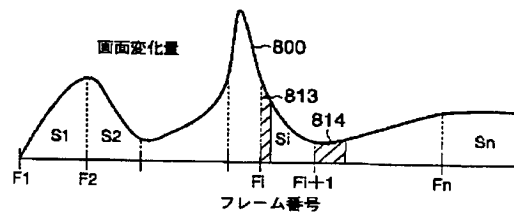
【図19】



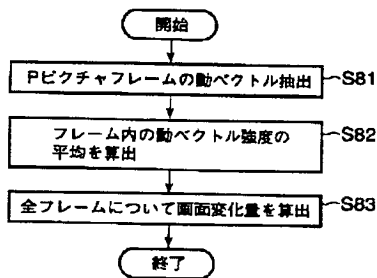
【図20】



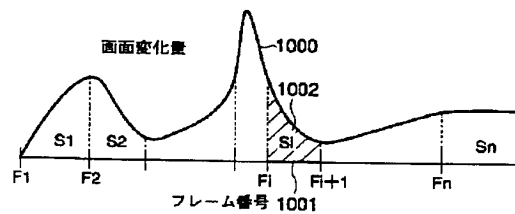
【図21】



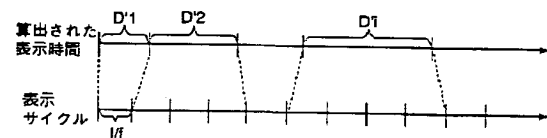
【図23】



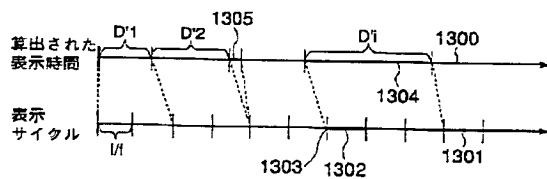
【図25】



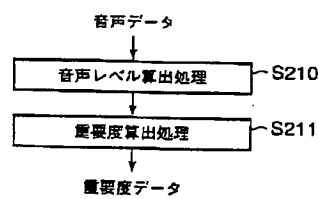
【図29】



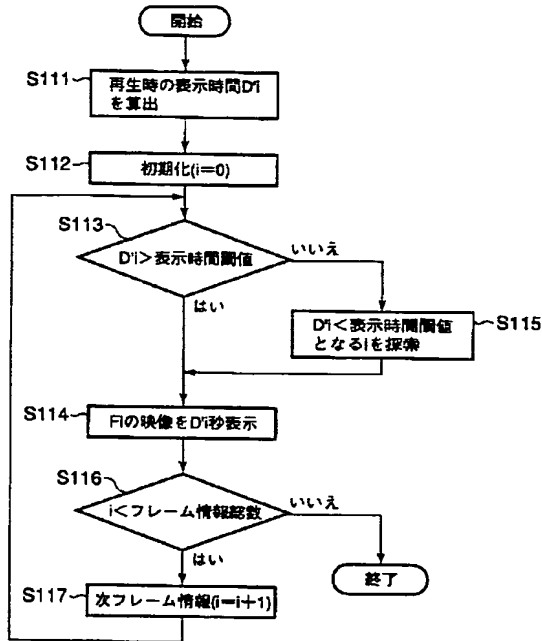
【図28】



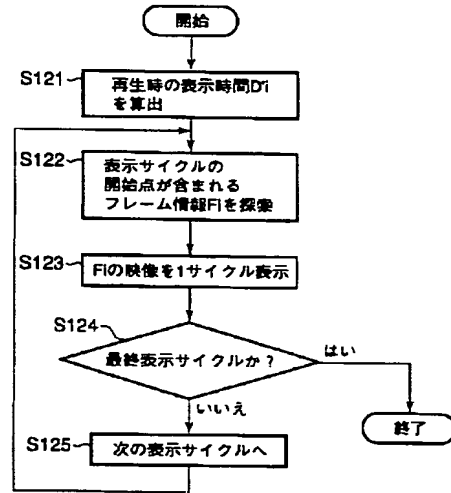
【図38】



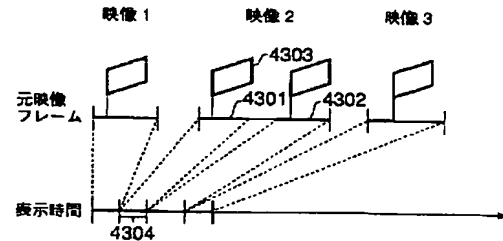
【図26】



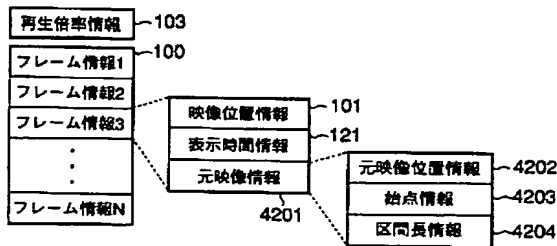
【図27】



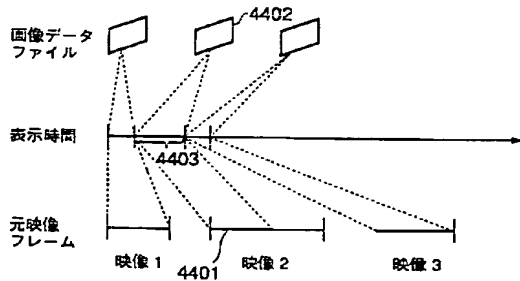
【図31】



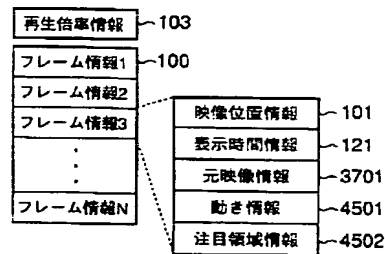
【図30】



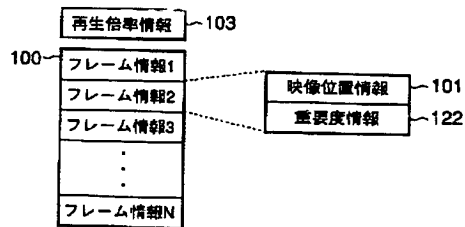
【図32】



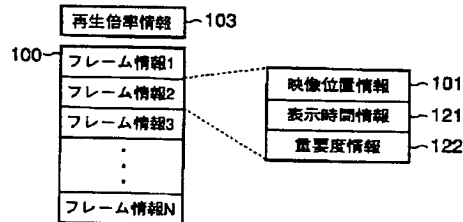
【図33】



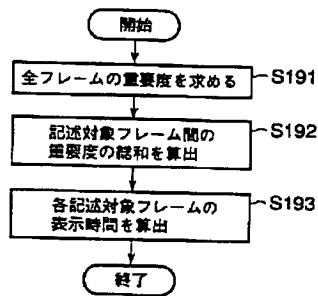
【図34】



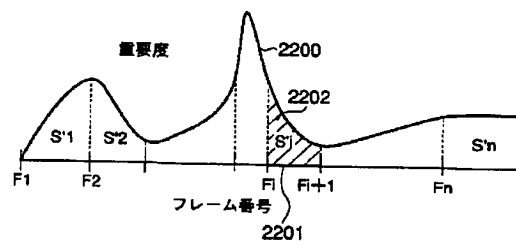
【図35】



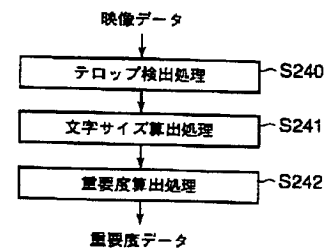
【図36】



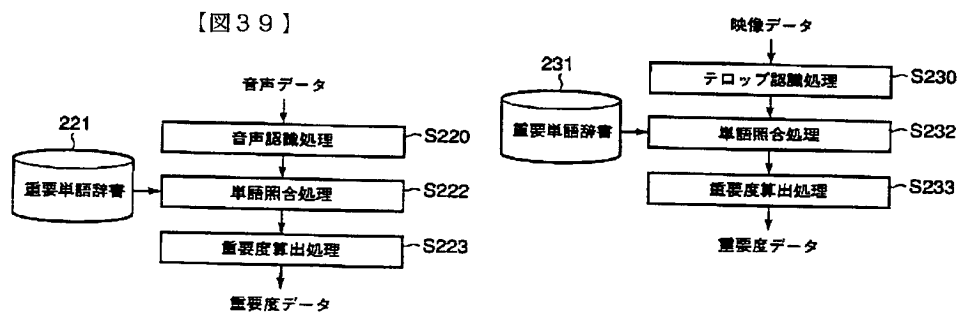
【図37】



【図41】



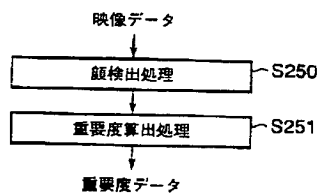
【図40】



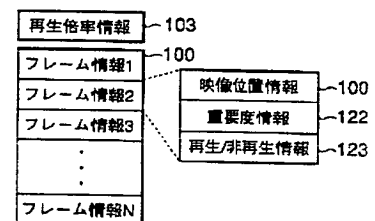
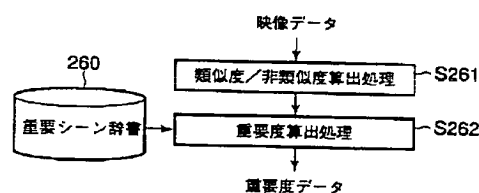
【図39】

【図45】

【図42】

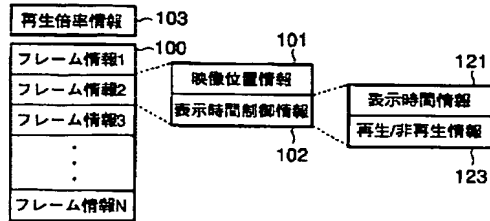


【図43】

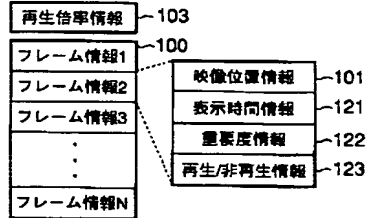




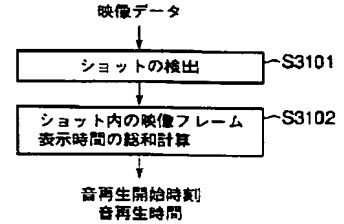
【図44】



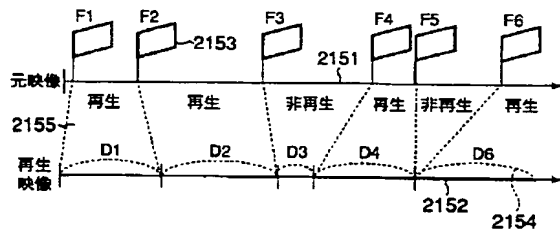
【図46】



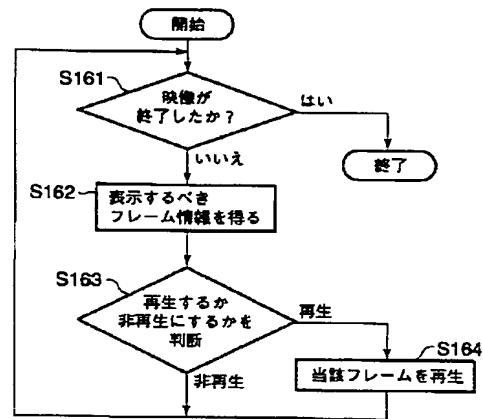
【図53】



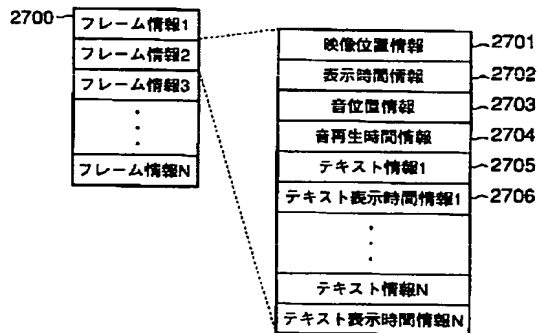
【図47】



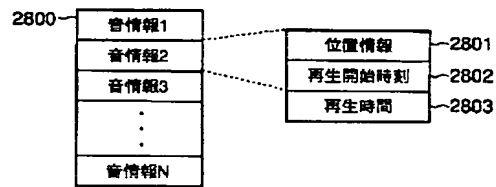
【図48】



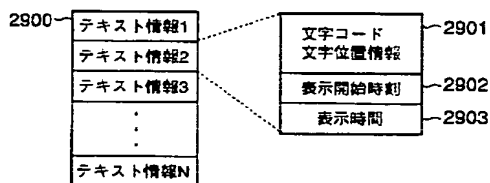
【図49】



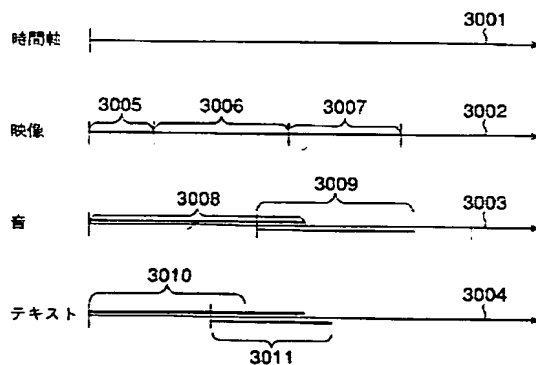
【図50】



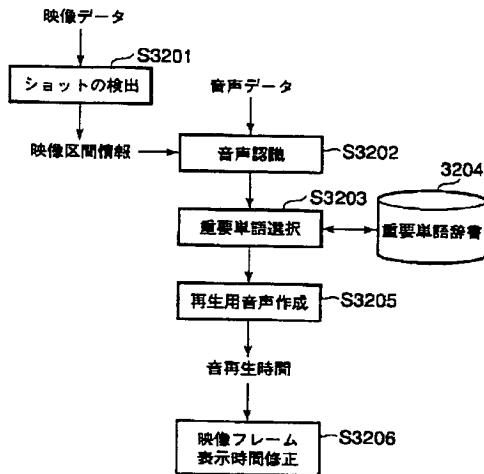
【図51】



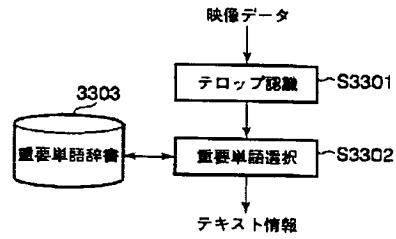
【図52】



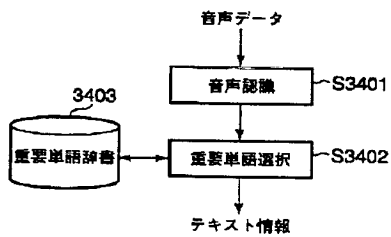
【図54】



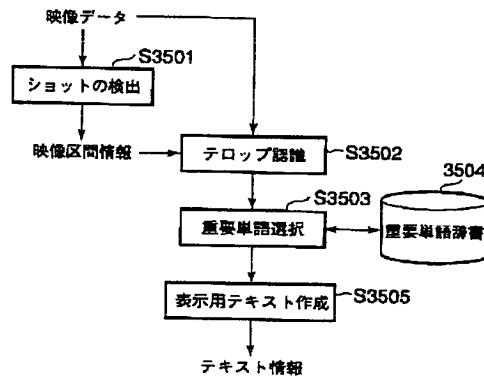
【図55】



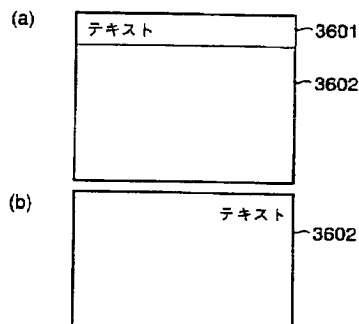
【図56】



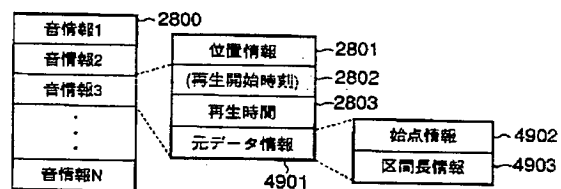
【図57】



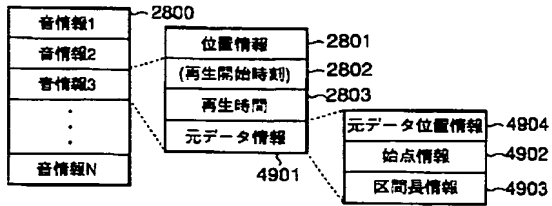
【図58】



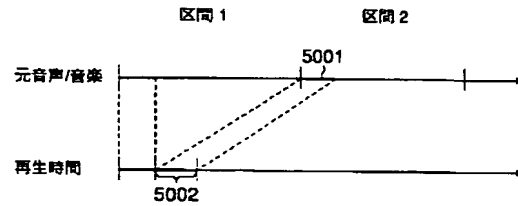
【図59】



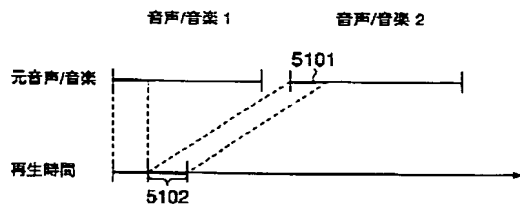
【図 60】



【図 61】



【図 62】



フロントページの続き

(72)発明者 三田 雄志  
神奈川県川崎市幸区小向東芝町1番地 株  
式会社東芝研究開発センター内  
(72)発明者 山本 晃司  
神奈川県川崎市幸区小向東芝町1番地 株  
式会社東芝研究開発センター内

(72)発明者 増倉 孝一  
神奈川県川崎市幸区小向東芝町1番地 株  
式会社東芝研究開発センター内  
Fターム(参考) 5C052 AA01 AC01 CC11 DD04  
5C053 FA23 GA11 GB37 HA21 JA24  
5D044 AB05 AB07 BC03 CC04 FG18  
FG23

THIS PAGE BLANK (USPTO)

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record.**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☒ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)